SAMPLING OPTIMIZATION FOR SOIL CARBON ASSESSMENT IN A COMPLEX AGROECOSYSTEM OF THE NORTHEASTERN UNITED STATES

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Sonam Rinchen Sherpa

May 2014

ABSTRACT

Low-cost accurate methods for estimating soil carbon (C) stocks are needed if terrestrial C offset markets are going to be implemented in the United States. Accurately measuring C stocks is often prohibitively expensive due to high spatial variability and analytical costs, therefore the development of cost–effective sampling designs and methods of inference are critical. We evaluated sampling optimization approaches for estimating soil C baseline levels for a dairy farm in Harford, NY with multiple land uses, including cultivation of silage corn, alfalfa hay, pasture, and forest. Three hundred and nineteen samples were collected in a spatially balanced design over a 232 hectare area to a depth of 30 cm. Secondary variables including soil type, elevation, slope, cropping history, and manure application rate were assessed for correlations with soil C and suitability for sampling stratification. Random, stratified, and systematic sampling arrangements at three sampling densities (n = 253, 160, 83), were compared to the full sampling grid (n = 319) using both design-based and model-based approaches for soil C assessment. Total soil C stocks for the sampling area were estimated by three different approaches: i) spatial mean (SM) where total C stocks are calculated by the area-weighted average of the mean C stocks for each landscape unit; ii) ordinary kriging where the sum of the predicted values for the interpolation grid are used to determine total C-stocks; iii) (SSURGO) where average C stocks are based on estimates from the Soil Survey Geographic database with total C stocks calculated from the area-weighted average for each soil map unit. The systematic sampling

arrangement was preferred over random or stratified arrangements because RMSE increased little with the reduced sample size, and the distribution of soil C stocks for the lowest sampling density closely resembled the full sampling grid. Landscape units defined by manure application rate explained the spatial variability of soil C-stocks better than any other categorical variable. Model-based approaches provided more reliable estimates for soil C stocks than design-based approaches. SM resulted in a higher RMSE than OK, 20.7 and 23.1 Mg ha^{-1} compared to 18.0 and 22.4 Mg ha^{-1}, respectively. Additionally, when the sampling density was reduced from 319 to 83, OK estimates fluctuated less than SM, with mean and total soil C stocks for the entire farm differing by 2% from that of the full sampling grid. Estimates of total C stocks to 30 cm for the entire 232 ha sampling area ranged from 16217–20049 Mg. Model-based approaches provided the most reliable estimates of soil C stocks. SSURGO based estimates consistently underestimated soil C stocks by 2.6-18.1 % compared to the full grid sampling, but given the low cost this approach may be of interest in some circumstances.

BIOGRAPHICAL SKETCH

Sonam Sherpa was born on August 16, 1983 in Pequannock, New Jersey to Judy and Topgyal Sherpa. He has two older brothers, Jayson Rome and Pema Sherpa. Sonam attended high school at the Windsor Academy in Morristown, New Jersey. Upon completion of high school, he spent one year travelling around India and Nepal, studying Tibetan culture and religion.

Having experienced and embraced his cultural heritage Sonam returned to the United States ready to begin his higher education. In August of 2003, Sonam enrolled in the Pre-Environmental Science and Forestry program at Alfred State College, State University of New York in Alfred, New York. He completed his Associate in Arts degree in the May of 2005, and by August 2005 he was continuing his education at Cornell University Ithaca, New York. In January 2007, Sonam completed his Bachelors of Science at Cornell with a dual degree in Natural Resources and Plant Sciences.

Upon completion of his undergraduate degree, Sonam continued to explore his passion for ecology and the environment by joining the Cornell Soil Health Team, where he was able to enhance his understanding of agroecological systems and the techniques and skills required to study them. After four years of applied agricultural research experience Sonam decided to further his education once again by joining the graduate field of Horticulture at Cornell University.

ACKNOWLEDGMENTS

First I would like to thank my committee chair, Dr. David Wolfe, for not only being a great mentor, but for making me feel welcomed and appreciated both personally and professionally. Your critical feedback and positive support has allowed me to develop into a better scientist. I would also like to thank my other committee member, Dr. Harold van Es, for his advice regarding spatial statistics and reflectance spectroscopy. Additionally, I would like to thank Dr. Peter Woodbury and Dr. Zia Ahmed, for the wealth of advice they provided about sampling design and spatial modeling. Many thanks to Robert Schindelbeck for all of the advice he provided me throughout the years, and for his unending willingness to help others.

Thank you to my family and friends, without your love and positive influences I would not be here today. Field and laboratory work could have been completed without the help of Jeff Beem-Miller, Ryan Devlin, Stefan Karkuff, Jonathan Comstock, Paul Thomas, and Jessica Leonard. Finally, thank you to the United States Department of Agriculture's National Institute for Food and Agriculture, for so graciously providing financial support for my research and studies.

TABLE OF CONTENTS

Chapter 1

A Review of Sampling Designs and Methods of Statistical Inference for Soil Carbon Assessment in Agroecosystems

Abstract

A full greenhouse gas accounting for agroecosystems requires estimation of baseline soil carbon (C) stocks, which can be prohibitively expensive due to high spatial variability. Increasing soil C stocks is of great interest to researches and policy makers, as promoting practices that sequester soil C will also reduce farm greenhouse gas emissions, improve soil fertility, reduce erosion, and increase resiliency to climate change. In this review we provide an overview of sampling approaches and methods of statistical inference for soil C assessment, with a primary focus on model-based (geostatistical) and design-based (classical) statistical techniques. While there is still debate on whether design-based or model-based approaches are most appropriate for soil C assessment, or whether stratified, random, or systematic sampling are most appropriate for developing semivariograms used in geostatistical modeling, it seems clear that incorporating the spatial structure of soil C in any analysis will improve results when sufficient spatial autocorrelation exists at the scale of the study site. The specific objectives of a sampling campaign and the inherent biophysical characteristics of a site will ultimately dictate whether design-based or model-based approaches are most appropriate.

1. Introduction

Soil has a tremendous potential for storing atmospheric carbon (C), with managed ecosystems having a storage capacity of 55 to 78 GT (Lal, 2004). It has been estimated that soil represents the world's largest stock of terrestrial C (1500Pg), nearly twice the amount of C that is stored in the atmosphere, fauna, and vegetation combined (Bartholomeus et al., 2008). The conversion of natural ecosystems to intensively managed agricultural lands has resulted in a global reduction of soil organic C (SOC) (Post and Kwon, 2000; West and Post, 2002; Ogle et al., 2005). Tillage reduces SOC through two primary mechanisms: by aerating the upper soil profile which stimulates microbial decomposition of plant residue, and by disrupting soil aggregates that physically protect SOC from microbial decomposition. Soil mediated processes such as microbial respiration associated with decomposition, C and nitrogen (N) mineralization, and denitrification, along with photosynthetic CO_2 uptake and conversion to C-rich biomass and root exudation by plants, play a major role in global greenhouse gas (GHG) fluxes, and atmospheric levels of GHGs. Recent studies by the United States Environmental Protection Agency (EPA) have estimated that changes in management by the agricultural and forestry sectors have the capacity to reduce the annual U.S GHG emissions by 10-25 percent (EPA 2005).

Improving agricultural practices by reducing tillage and planting cover crops can sequester C at a rate of 0.1 to 1.3 Mg C ha^{-1} yr^{-1} (Lal, 2004). While this alone would have only a modest effect on restoring the global C balance, increasing SOC often comes at little cost or even increases profits for the farmer. In addition to contributing to international goals of mitigating global climate change through reduction of atmospheric GHGs, increasing SOC has potential co-benefits, such as improving crop production in areas with degraded soils, and building resilience

to climate change as soil organic matter improves water holding capacity, infiltration and drainage, cation exchange capacity, biological activity, and nutrient cycling. Thus, SOC sequestration is a promising win-win strategy for large scale GHG reduction. However, despite the many benefits associated with increasing SOC, conventional agricultural practices that reduce SOC, such as fallowing and intensive tillage via moldboard plowing, remain common in agroecosystems throughout the United States.

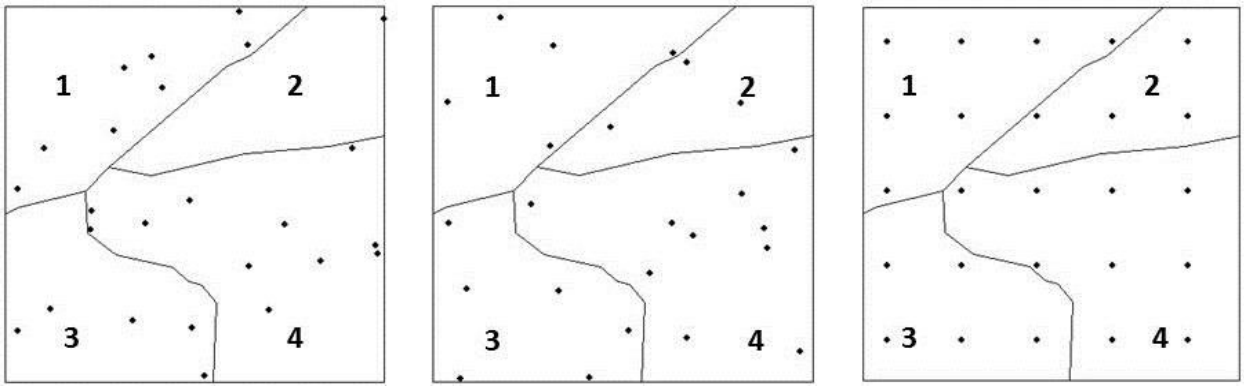
Market based C trading has been proposed as a mechanism for incentivizing agricultural management practices that store C in the soil (Pautsch et al., 2001; Antle et al., 2002; Conant et al., 2011). Under a market based trading system farmers would be paid for the amount of soil C stored due to some change in soil management. While adoption of specific management practices is known to affect soil C, market based trading for soil C is likely to require that payments be based on the actual quantity of C stored (Antle et al., 2003). Soil C is typically extremely variable at both large and small spatial scales, thus accurately determining soil C levels over a large area is prohibitively challenging due to high sampling and analytical costs (Conant et al., 2003). Additionally, for accountability, a C market system requires soil C levels to be reported on an area basis to a specified depth (i.e. volumetric basis) not just as C concentration. A C stock inventory typically requires soil sampling with an auger of specific area to a specific depth and accurate estimation of the bulk density (e.g., g soil/cm^3). Soil variability, soils with a tendency for shrinking and swelling, soils that are high in coarse rock fragments or organic matter, and comparing fields that differ in tillage pose many challenges for measuring bulk density and estimating C stocks in a cost effective manner.

Developing low cost approaches for evaluating SOC baselines and changes in SOC stock, from the field to the global scale has been identified as a high priority by researchers and policy-

makers (Conant et al., 2011). To date, relatively few studies have focused specifically on sampling optimization for soil C assessment at the agroecosystem scale across multiple land uses, and even fewer studies have evaluated how design-based inference compares to model-based inference with varying soil sampling optimization approaches (Brus and De Gruijter, 1997; Pringle et al., 2011). In this review we will examine current approaches for soil C assessment, with a primary focus on options regarding spatial soil sampling, and statistical and modeling methodology.

2. Spatial Soil Sampling

While there are many soil sampling approaches, we will limit our discussion to three basic categories (Figure 1.):



(a) **(b)** **(c)**

Fig. 1. The locations of 25 sampling points on a soil series map using (a) simple random sampling , (b)stratified random sampling, and (c) systematic sampling designs. For the random sampling design soil series 2 is under sampled. For simple and stratified random sampling, sample clustering evident. For systematic sampling samples are equally spaced throughout the area.

2.1. Simple Random Sampling (Fig. 1.a): sample locations are chosen completely at random.

One of the main attractions of simple random sampling is that it is unbiased and analysis of the data is relatively straight-forward, allowing for the use of classical statistical methods such as analysis of variance (ANOVA). The drawback of simple random sampling is that the sampling variance is usually larger than with other sampling designs, resulting in a need for more samples and increasing sampling costs. This inefficiency is due to the fact that samples may provide uneven spatial coverage such that large areas are under-sampled and over-sampled areas provide redundant information. Additionally, if individual estimates are needed for sub-areas within a study region, simple random sampling may not be able to provide this information.

2.2. Stratified Random Sampling (Fig. 1.b.): sampling region is divided into homogenous groups or "strata", and simple random sampling is conducted within each group.

Stratified random sampling addresses many of the drawbacks of simple random sampling by distributing samples across an area based on our knowledge of processes that influence soil C. Using this approach an area is divided into homogeneous groupings, were the weighted mean for the total area and the number of samples allocated to each strata is dependent on the respective area of each strata. Strata may be defined by the geographic area of ancillary variables such as elevation, land use, or soil type. By choosing homogenous strata one may reduce the sampling variance and improve efficiency. However, if variables used for stratification are not correlated with soil C or if strata are poorly defined, the sampling variance may increase (de Gruijter, 2006). Additionally, if stratified random sampling results in severe sample clustering and poor spatial coverage, empirical variograms and model-based inference may be unreliable.

2.3 Systematic Sampling (Fig. 1.c): samples are located on a regular grid.

Systematic sampling is well suited for model-based inference (see 4.0, below), where randomness is introduced by the model rather than the sample design. Systematic sampling is simple to implement, provides even spatial coverage, and often allows for accurate estimates of the empirical variogram. While systematic sampling requires points to be equally spaced throughout the study area, if the spacing between points is too large relative to the study area some strata may be under-sampled. Additionally, if a cyclic or directional pattern in spatial variance is present, i.e. anisotropy, systematic sampling may produce erroneous model estimates. Other important issues to consider are the pattern and orientation of the sample grid. Square, triangular, and hexagonal are the three most common grid patterns, An equilateral triangle design has been shown to provide the most reliable estimate of semivariograms as the average distance to unsampled locations is minimized, and anisotropy is better detected (Yfantis et al. 1987; McBratney et al. 1981).

All three of the above sample designs may be used for model-based or design-based inference; however, systematic sampling should only be used for design-based inference if the initial grid position was chosen at random so that assumptions about the selection probabilities of sampling locations are not violated.

2.4. Soil sampling in relation to planned approach for statistical inference.

The soil sampling approach and planned methods of statistical inference are interdependent: in order to gather the most appropriate spatial soil data, one must understand and consider the methods of inference most appropriate to empirical questions being asked. Design-based sampling, often used for classical survey sampling with maximum interest on the mean

value, regards the population of values in a region as fixed, and sampling locations are randomly assigned. Furthermore, when calculating weighted averages, weights are based on selection probabilities rather than the geographic location of samples. In model-based sampling, values of a region are thought of as just one realization of a stochastic model, and weights are based on the covariance between observations as determined by the geographic location of the observations. In the model-based approach, sampling locations need not be random because randomness has been introduced through the model (de Gruijter et al., 2006).

The benefits of design-based sampling are maximized if global quantities such as the spatial mean are of interest, while model-based sampling becomes more appropriate as the number of sub regions being estimated (i.e. spatial resolution) increases. Even when estimating global quantities over large areas, model-based approaches may be more appropriate if the variable of interest exhibits strong spatial autocorrelation at the scale of the study area, and if the sample size is large enough to benefit from the autocorrelation (De Gruijter et al., 2006). Autocorrelation is described in more detail in the following sections.

Differences between the soil sampling approaches for design- vs. model-based sampling can be demonstrated by an example from Brus & de Gruijter, (1997), where an area is repeatedly sampled at 25 locations and a 0/1 indicator variable *z* is measured to determine the proportion of the area with a value 1 (Figure 2.). In the design-based approach 1a, 1b, and 1c, values of *z* have fixed locations while the locations of samples vary randomly. In the model-based approach 1a, 1d, and 1e, sample locations are fixed, and values of *z* vary according to the realizations of a stochastic model. As a result of this difference, statistical inference from model-based approaches is based on a stochastic model describing real world variation, while design-based

approaches rely on “the selection probabilities of sampling locations as determined by the random sampling design” (De Gruijter et al., 2006).

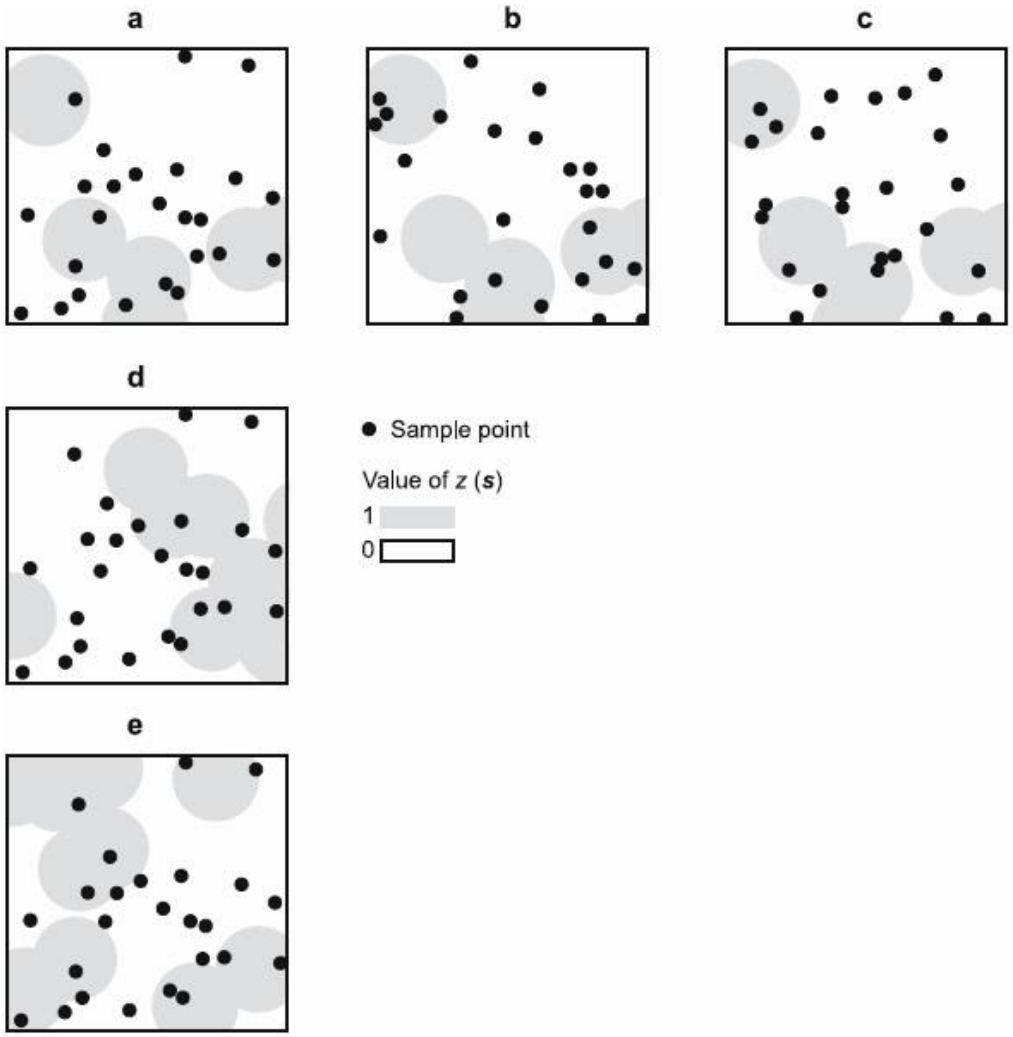


Fig. 2. Design-based (a,b,c) and model-based (a,d,e) sampling approaches for repeated sampling. In the design-based approach values of z have fixed locations and sample points have random locations. In the model-based approach values of z have random locations and sample points. (from Brus and De Gruijter, 1997)

3. Design-Based Inference

Design-based approaches are well suited for soil C assessment when the sole objective is to determine the spatial mean i.e. the amount of C stored in some region to a given depth, rather

than how C is distributed throughout the region. A simple random sampling design ensures that all sampling locations are selected with an equal probability, thus it is assumed that the sample mean, sample variance, and variance of the sample mean [μ_s, σ^2_s, and , $\sigma^2_s(\mu_s)$], provide unbiased estimates of the population (by Allen et al., 2010):

$$\mu_s = \frac{1}{n}\sum_{i=1}^{n} y_i \quad (1)$$

$$\sigma_s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \mu_s)^2 \quad (2)$$

$$\sigma_s^2(\mu_s) = \frac{\sigma_s^2}{n} \quad (3)$$

Stratified random sampling will also provide unbiased estimates of μ, σ^2_s, and, $\sigma^2_s(\mu_s)$; however the formulas must be adjusted to account for the change in section probabilities resulting from stratification (by Allen et al., 2010):

$$\mu_{st} = \sum_{h=1}^{H} a_h \mu_h \quad (4)$$

$$\sigma_{st}^2 = \left(\frac{1}{n}\sum_{i=1}^{n} y_i^2\right) - \mu_{st}^2 + \sigma_{st}^2(\mu_{st}) \quad (5)$$

$$\sigma_{st}^2(\mu_{st}) = \sum_{h=1}^{H} a_h^2 \sigma_h^2(\mu_h) \quad (6)$$

Where H is the number of strata; a_h is the proportion of the area representing the hth stratum; and μ_h is its mean.

4. Geostatistics and Model-Based Inference

4.1. Spatial autocorrelation and variogram estimation

Statisticians have been quantifying spatial variability since the early days of classical statistics when Mercer and Hall (1911) observed spatial patterns in crop yield between small plots at the Rothamsted Experimental Station. They noticed that barley yields in adjacent plots were more similar than those that were far apart, and attributed this to two different sources of variability: one that was autocorrelated and another that was completely random. When R. A. Fisher came to Rothamsted during the 1920's he came up with concept of blocked experimental designs to separate these sources of variance by plot layout and analysis of variance (ANOVA) procedures, which set the stage for design-based sampling and inference (Webster and Oliver., 2001).

Today, geostatistical model-based approaches, rely on the modern concept of spatial autocorrelation, first introduced by Matheron (1962), to develop methods for predicting values at un-sampled locations in space. The underlying assumption is samples located closer in space will be more similar than those that are further away. The degree of spatial autocorrelation can be quantified by the experimental semi-variogram (Figure. 3):

$$\bar{\gamma}(\mathbf{h}) = \frac{1}{2n(\mathbf{h})} \sum_{i=1}^{n(\mathbf{h})} \{z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})\}^2 \quad (7)$$

Where $\bar{\gamma}(\mathbf{h})$ is the average semivariance as a function of **h**; $n(\mathbf{h})$ is the number of point pairs as a function of **h**; $z(\mathbf{x}_i)$ is the observed value of z at the ith location; and $z(\mathbf{x}_i + \mathbf{h})$ is the observed value of z at a distance of **h** from $z(\mathbf{x}_i)$. The empirical variogram estimates the variance

at increasing intervals of distance; however, it is often quite noisy and must be modeled by a theoretical function (Webster, R Oliver, M. A., 2001; Allen et al., 2010). Spherical, exponential, and Gaussian are three of the most commonly used functions for modeling the semi-variance. The spherical model is often most appropriate for estimating the theoretical semi-variogram for soil C. The spatial variability of a site is effectively described by three components of the semi-variogram: the range, sill, and nugget (Figure. 3). The range can be thought of as the zone of influence, or the distance at which samples are no longer spatially correlated. The sill represents the total observed variance, and the nugget represents the portion of the variance that cannot be explained by spatial correlation. The nugget is often the result of measurement error and spatial trends that occur at a scale smaller than the minimal sampling distance (Cressie, 1991).

Soil sampling to quantify spatial autorcorrelation and develop a semi-variance function is usually an essential first step in geostatistical, model-based inference for spatial data. As indicated previously, unlike the case for design-based inference, model-based inference does not require randomization of soil sampling to develop the semi-variogram. Variograms are most commonly estimated using Matheron's (Matheron, 1962) methods of moments variogram (MoM) equation, but sample sizes of 100 to 150 are required to adequately estimate the variogram (Webster and Oliver, 1992). This relatively large initial sample requirement has led some soil scientists to stray away from model-based approaches to soil C assessment. Various authors have shown that the range of spatial autocorrelation can be used to determine the sampling interval so that redundant information is minimized (Kerry and Oliver, 2003).

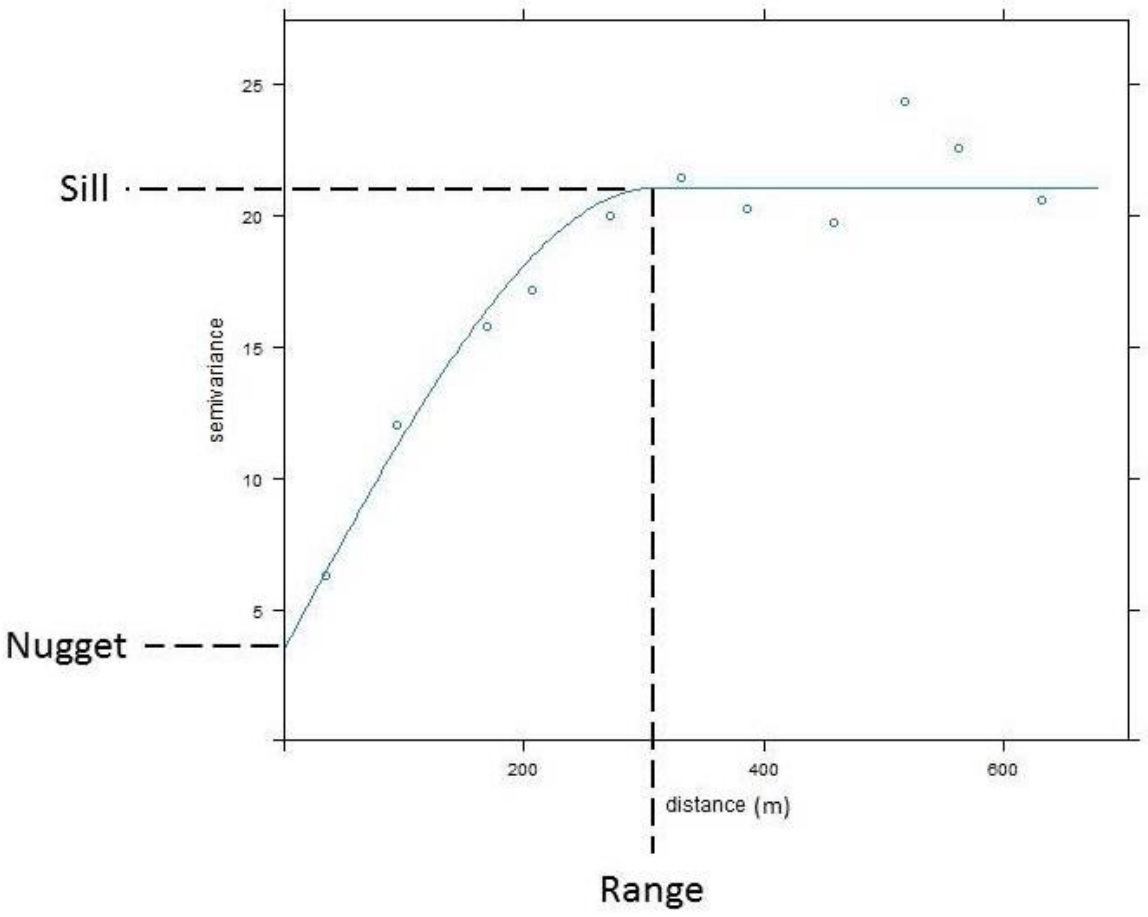


Fig. 3. Empirical and fitted variogram for soil clay content. Experimental values are plotted as points, solid line is the fitted variogram using a spherical model. The sill, range, and nugget are labeled as such.

In cases where the range is large relative to the size of a field or farm, the required sample size may not be large enough to accurately estimate the variogram. Computing variograms by a residual maximum likelihood (REML) approach has been proposed as a solution to the large sample requirement of MoM variograms (Lark, 2000).

Kerry and Oliver, (2007) compared MoM and REML variograms for soil clay content at four field sites in England. Performance of the two different methods was evaluated based on

sample size and arrangement. It was found that predictions based on REML variograms were generally more accurate than those of MoM variograms when there were fewer than 100 sampling sites. They suggest that a sample size of around 50 samples with an appropriate spacing is adequate for kriging soil properties, for both purposes of precision agriculture and detection of soil contaminates. Computing variograms by REML seems like an obvious choice for soil C assessment as sampling budgets are often constrained. The lack of wide spread application of REML variograms can be attributed to 3 main causes:

1) Available software for estimating REML variograms is not user friendly.

2) The REML method assumes the data follow a multivariate Gaussian distribution, which is impossible to verify; at best we can ensure that the data approximate a normal distribution.

3) The REML approach is computationally intensive, and until recently it was believed that REML computations had an upper limit of 150 samples, though Bellamy et al., (2005) showed that sample sizes of up to 1000 are feasible.

4.2. Geostatistcal models

The focus of the discussion here will be limited to two specific model-based methods of inference: ordinary kriging (OK) and regression kriging (RK) which have been shown to provide reliable estimates for soil C assessment.

4.2.1. Ordinary kriging

Kriging is a group of interpolation methods named after the South African mining engineer D. G. Krige, who during the 1950's came up with an empirical method for estimating the true ore grade distribution of gold and other minerals from a sampled distribution (Cressie, 1991). Matheron (1962) refined Krige's original work to develop the "multiple regression procedure for arriving at the best linear unbiased [predictor]" that we now know as kriging (Cressie, 1990). The purpose of kriging is to estimate the value of a random variable at unsampled locations, where predictions consist of the weighted average of nearby samples, and weights are assigned based on patterns in spatial variability as determined by the semivariogram model. Ordinary kriging is one simplest and most commonly used forms of kriging, and is given by the formula (Hengl et al., 2007):

$$\hat{z}(s_0) = \sum_{i=1}^{n} \lambda_i \cdot z(s_i) \qquad (8)$$

Where $\hat{z}(s_0)$ is the OK predicted value z at location (s_0), n is the number of samples in the search neighborhood, and λ_i is the weight assigned to the ith observation $z(s_i)$. Kriging is an optimal prediction method in that weights are assigned to sample points within a search neighborhood in such a way that the estimation or kriging variance $E[\{\hat{z}(s_0) - z(s_0)^2\}]$, is minimized, and that estimates are unbiased (Webster and Oliver, 2001).

4.2.2 Regression kriging

The increased computational power of computers, advancements in information technology, and the ready availability of massive environmental databases (e.g., Soil Survey Geographic Database (SSURGO), digital elevation models, and satellite imagery) have all

contributed to the recent surge of research focusing on predictive modeling and mapping of soil properties (McBratney et al., 2003). With current international concerns regarding food security and climate change mitigation, it is no surprise that a substantial amount of the above mentioned work has concentrated on quantifying soil C levels. Many of the more promising approaches for predicting SOC rely on a combination of observations from the area of interest, and auxiliary information on soil forming processes such as climate, topography, and vegetation. These latter data are particularly useful as they are often available at little or no cost, provide relatively high resolution information over large spatial extents, and perform well as model covariates (Minasny et al., 2013).

Auxiliary data is often readily available or inexpensive to measure, and may provide valuable information about the main variable of interest. RK is a hybridized geostatistical approach that builds on OK in the sense that it combines regression using auxiliary variables with OK. During RK soil C is first predicted from auxiliary variables using a regression model, then the residuals of the regression are interpolated using OK (Hengl et al., 2004). RK has been shown to improve estimates of soil C and provide better predictions than either regression analysis or OK would on its own (Knotters et al., 1995; Bishop and McBratney, 2001; Simbahan and Dobermann, 2006; Hengl et al., 2007). Auxiliary variables that have been used for prediction of soil C by RK include soil series, elevation, terrain indices, electrical conductivity (EC), surface reflectance, crop yield, land use, precipitation, reflectance spectroscopy, and others (Simbahan et al., 2006; Kravchenko and Robertson, 2007; Bilgili et al., 2011; Zhang et al., 2012; Minasny et al., 2013; Cambule et al., 2013). The equation for RK is given by (Hengl et al., 2004):

$$\hat{z}(s_0) = \sum_{k=0}^{p} \hat{\beta}_k \cdot q_k(s_0) + \sum_{i=1}^{n} w_i(s_0) \cdot e(s_i); \quad (9)$$
$$q_0(s_0) = 1;$$

Where$\hat{z}(s_0)$ is the RK predicted value z at location (s_0), β_k are the regression coefficients, λ_i are the kriging weights determined by the variogram of the regression residuals $e(s_i)$ at location s_i, $q_k(s_0)$ is the kth auxiliary predictor at location (s_0), and p is the number of predictors.

One of the main drawbacks of RK over OK is the additional computational steps required to perform this method. Hengle et al., 2007 provide a methodological framework for implementing RK by means of the following eight steps.

1) Define the regression model as predicted from auxiliary variables. Stepwise regression may be used to eliminate unnecessary variables.
2) Compute the ordinary least squares (OLS) residuals from the regression for all sample locations.
3) Model the variogram of the OLS residuals.
4) Use the modeled variogram of the OLS residuals to estimate the generalized least squares (GLS) coefficients. This ensures that the regression accounts for spatial correlation of the residuals.
5) Compute the GLS residuals for all sample locations.
6) Model the variogram of the GLS residuals. Unless significant clustering of sample points occurs, the variogram of the GLS and OLS residuals will be quite similar.
7) Interpolate the GLS residuals using OK with the GLS variogram.
8) Add the GLS trend surface to the interpolated GLS residuals at all prediction locations.

4.2.3 Model evaluation

It is essential to assess the performance of any method of inference, yet many published studies on modeling of soil C provide little or no information on validation methods or prediction accuracy (Minasny et al., 2013). Model performance is commonly evaluated by the goodness of fit and prediction error. The coefficient of determination (R^2) describes the model's goodness of fit, while the root mean squared error of the prediction (RMSE) informs us about prediction accuracy.

$$\mathrm{RMSE} = \sqrt{\frac{1}{l} \cdot \sum_{j=1}^{l} [\hat{z}(s_j) - z^*(s_j)]^2} \quad (10)$$

RMSE is calculated by comparing the predicted value $\hat{z}(s_j)$ at a validation sample location with an actual observation $z^*(s_j)$, where l is the number of validation samples. Several commonly used methods for validation include: leave-one-out-cross-validation (LOOC), internal validation, or independent sampling for validation. In the LOOC approach the model is calibrated with one sample removed, this sample will then be used to validate the model. The removed sample is then returned to calibration set, and the entire procedure is repeated until all samples have been removed. Internal validation follows the same basic procedure except that the sample data are divided into separate calibration and validation sets. Typically 70% of samples are used for calibration while the remaining 30% are used for validation; designation of sample use may purposive or random. Independent sampling for validation will provide unbiased and valid estimates of model performance, and is preferred over other forms of validation (Brus et al.,

2011). Brus *et al.*, 2010, found that LOOC is the best option when independent validation is not possible. In order for a validation sample to be truly independent from calibration samples some form of probability sampling is required. Independent sampling for validation will provide superior estimates of model performance; however, due to the costs and challenges associated with additional random sampling many researchers prefer LOOC or internal validation.

5. Applications of Design- and Model-Based Approaches to Soil C Assessment

5.1. Soil sampling optimization

Approaches for sampling optimization will vary considerably depending on the type of inference being used, but the overall goal is the same–maximize the quality of an estimate while minimizing its cost (i.e., sample size required). Quality is defined as some statistical measure of accuracy, precision, or reliability. Common measures of quality include RMSE, standard error, error variance, and half-width of confidence intervals (De Gruijter, 2006). For model-based approaches, one attempts to find the best sampling pattern and smallest sample size for some threshold measure of quality. In the design-based approach, the sample design itself is regarded as stochastic and thus cannot be optimized; instead optimization focuses on the selection of the design type and the randomization restrictions that are applied. Stratified random sampling is an example of a randomization restriction, where a given number of samples are restricted to a certain number of strata. Optimization is focused on the shape, size, and location of strata and the number of samples within each stratum.

5.2 Design-based inference for soil C assessment

The purpose of soil C inventory in agroecosystems is to determine how much soil C has been stored or lost on a farm as the result of some management practices. This is essentially the same as estimation of the global mean, for which design-based inference is best suited. Following this logic several authors have suggested that stratified random sampling and design-based inference is optimal for soil C inventory(McKenzie et al., (2000); Lark, 2009; Allen et al., 2010; Singh et al., 2013). In a recent review on sampling designs for measuring soil C in Australian grazing lands Allen *et al*., (2010) makes the case that design-based sampling and inference is more efficient than geostatistical (i.e., model-based) approaches for measurement of soil C. Similarly, McKenzie *et al*. (2000) proposed a stratified random sampling approach with a minimum of 4 replicates per strata for soil C accounting in Australia at the national scale. Singh *et al*., (2013) developed a framework for quantifying sampling costs and attempted to determine the most efficient strategy for estimating soil C stocks at the field scale. They compared different sampling arrangements, quadrat sizes, and random or stratified placement of the quadrats. It was found that quadrats with a 100 m radius were most efficient, and that stratified placement of quadrats reduced standard errors.

Lark, (2009) raised the issue that sampling requirements differ for soil C baseline inventory vs. monitoring soil C *change* over time because variability of soil C status and soil C change may not be the same. Information on variability of change is usually not available, and it is unsafe to assume that sample requirements for detecting soil C change can easily be computed from inventory data alone. Additionally, the sample size for determining soil C baseline values with a given level of confidence may not be the same as the required sample size for detecting change with the same level of confidence. When monitoring soil C change, Lark recommends

revisiting the same sample location(s) rather than repeated random sampling, as spatial variability is minimized and required sample size is reduced. Lark recommended using stratified random sampling to establish baselines with paired re-sampling at common sites. Using this method it was found that determining soil C change due to changes in land use required fewer samples than did the initial survey for estimating baseline levels.

5.3. Model-based inference for soil C assessment

In order to fully realize the benefits and drawbacks of model-based inference we must first determine the optimal sample pattern and number for a given method of inference. To address this question, Simbahan and Dobermann (2006) developed sampling optimization approaches for model-based inference and evaluated performance for soil C prediction to a depth of 30cm in three crop production fields. Three optimization approaches were considered 1) optimize for the generation of a variogram defined *a priori*, 2) spread sample points evenly over the whole region, 3) a combined approach where a portion of samples are allocated for each of the first two approaches. Cluster analysis based on constrained spatial simulated annealing was used to define strata from several auxiliary variables, and stratified random sampling was performed in addition to the three optimization sampling approaches. Auxiliary variables used in spatial classification were EC, digital elevation model (DEM), slope, surface reflectance, and soil series. Sampling approaches were compared using four sampling densities (50,100, 150, 200 samples) with either OK or RK. The combined optimization approach produced the best results, which were slightly better than stratified random sampling, with RK producing the smallest error for all designs. Simbahan and Dobermann (2006) found that using auxiliary information for sampling design and local prediction can greatly reduce sampling costs and increase prediction accuracy. A minimum of 1.5 -2 samples ha^{-2} was required for the 68 ha area.

Similarly, Kerry and Oliver (2003) examined the potential of using variograms from auxiliary variables (yield, EC, and aerial photographs) to optimize sampling for soil properties. They found that the scale of spatial variation obtained from aerial photographs was similar to that of soil properties and could be used to determine the sampling interval for soil surveys. Bilgili *et al*. (2011) compared the performance of OK, RK, and co-kriging (COK) using simple random and systematic sampling for predicting soil organic matter across various sample sizes. They found that COK out-performed OK and RK for nearly all sample sizes, and random sampling generally produced better results.

5.4. Using the Soil Survey Geographic database for soil C assessment

Both design-based and model-based approaches for soil C assessment often require a sampling budget that can be prohibitively expensive for even modestly sized areas. As such some have looked towards publicly available soil surveys to improve sampling efficiency and model predictions (Simbahan and Dobermann, 2006; Minasny et al., 2013). The Soil Survey Geographic (SSURGO) database provides detailed spatially referenced soil data for most regions of the United States, including percent soil organic matter (which can be used to estimate soil C concentration), or actual soil C concentration measurements. SSURGO data comes at no cost to users, provides valuable information about the spatial distribution of soil C, and has even been used to estimate C stocks at the U.S. national scale (West et al., 2010). Gelder et al., (2011) showed that soil C concentration in top 15 cm could be predicted at three field sites in Iowa by linear regression using SSURGO estimated C and surface reflectance from aerial imagery as predictors. Model R^2 and RMSE values ranged from 0.60 to 0.82 and 0.35 to 0.76%. SSURGO shows promise as a tool for estimating C stocks. This potential needs to be further investigated for different regions, soil types, and cropping systems.

6. Summary

When considering optimization approaches for model-based inference it is important to consider how different sampling methods for computing variograms will influence model choices. Various authors have favored either random or fixed sampling for estimating C-stocks, yet relatively few studies on sampling optimization have compared both methods using design-based and model-based approaches. Pringle *et al*., (2011) examined optimal sampling schemes for estimating soil C stocks for cattle grazing lands in Australia. Design-based sampling approaches were most appropriate as spatial autocorrelation was weak at this site. Systematic sampling was more statistically efficient than either simple or stratified random sampling, but there was less difference between designs as sampling area increased. Systematic sampling is well paired with model-based inference. However, when applied to design-based inference, the estimated variance will need to be approximated, leading to an inflated estimate. Pringle argued that stratified random sampling combined with design-based inference was optimal for estimating soil C stocks, particularly when spatial autocorrelation is weak at a site. They suggested an allocation of 25 samples per land unit using a stratified random sampling approach with stratification based on soil type and land use, and a minimum of 2 samples per stratum.

Mooney *et al*. (2007) found that spatial autocorrelation could be used to reduce standard errors and narrow confidence intervals for sample estimates of the mean quantity of soil C sequestered per hectare within a region. This reduction in uncertainty could reduce transaction costs associated contracts for C-credits, which would in turn increase C offset payments to producers. Benefits from accounting for spatial autocorrelation were not uniform across all strata or regions. The degree by which strata or regions will benefit from accounting for spatial

structure should be empirically verified. The authors recommended stratified random sampling for soil C assessment.

A similar approach was taken by Worsham *et al*., (2010) where they evaluated how land cover affects the spatial structure of soil C in forests in Georgia, USA, and how this will in turn influence prediction error and future sampling. They found that the standard error for mean soil C could be reduced when spatial autocorrelation was accounted for, and sample size could be reduced without a loss of accuracy. However, in contrast to Mooney et al. (2007), Worsham et al. (2010) found that model-based estimates for soil C using a systematic sampling protocol was more efficient than using a randomized approach.

Xiao *et al*., (2005) compared the efficiency of design-based and model-based approaches in terms of cost and accuracy for estimating the local and global mean vegetation cover in Fort Hood, Texas. They found that model-based approaches were more cost-efficient than classical designs because redundant information is reduced when spatial dependence is accounted for in the sample design. Kravchenko *et al*., (2006) showed that spatial structure could be used to improve the statistical efficiency of randomized complete block design (RCBD) experiments, and that soil C change between treatments could be detected sooner with a smaller sampling budget when spatial structure was accounted for. Additionally, they found that the minimum difference to detect for a given sample size decreased as the range on the variogram increased.

7. Next Steps

While there is still debate on whether design-based or model-based approaches are most appropriate for soil C assessment, or whether stratified random or systematic sampling are most appropriate for developing semivariograms used in geostatistical modeling, it seems clear that

incorporating the spatial structure of soil C in any analysis will improve results when sufficient spatial autocorrelation exists at the scale of the study site. The specific objectives of a sampling campaign and the inherent biophysical characteristics of a site will ultimately dictate whether design-based or model-based approaches are most appropriate. However, addressing the following issues will allow us to make more well-informed decisions of how best to sample for and estimate soil C stocks:

1. There is a need for more studies comparing the costs and benefits of design-based versus model-based sampling and inference, so that we may make better informed decisions regarding optimized approaches for soil C assessment. Are the gains in efficiency from utilizing auxiliary variables and model-based inference worth the costs of the specialized personal and time required for compiling and analyzing the additional information? Furthermore, there is need for improved methods for quantifying costs associated with soil C assessment. A systematic framework for assessing and comparing sampling, analytical, and data analysis costs must be developed before we can fully compare the efficiency of different sampling approaches.
2. Before model-based approaches are widely adopted for soil C assessment, minimum sample size requirements must be less than 100 to 150 samples. The potential of computing variograms by REML with fewer samples needs to be further investigated, and studies comparing design-based and model-based approaches must include REML variograms as well as MoM in their analyses.
3. *A priori* knowledge regarding the spatial structure of soil C will greatly improve sampling efficiency because the distance between sampling points can be set to reduce redundant information. Future studies should evaluate the feasibility of estimating

variograms for soil C from auxiliary variables so that informed decisions about sampling requirements can be made without reconnaissance sampling.

4. Previous research on soil C assessment in agroecosystems has mostly focused on global/regional or plot/field scales; this is due to the fact that government agencies and policy makers make management decisions on large regional scales while the basic unit of management for farmers is at the field scale. If our objective is to measure C stocks at the farm scale it may be beneficial to view the entire farm as a single unit rather than a sum of its parts. More research is needed to determine the ideal scale of measurement for quantifying soil C stocks within a farm with multiple land use and vegetation types (i.e., at the agroecosystem scale).
5. Surprisingly few studies have evaluated optimization approaches for soil C assessment in the NE United States. In this region a large proportion of corn-based agroecosystems are coupled with livestock operations, with much of the corn grown in NY consumed on farm for animal feed. Agroecosystems in this region are somewhat unique when compared to other regions of the U.S., as farms are often defined by diverse soil types and topography, inconsistent manure application rates, and a wide range of land uses including cultivation of silage, grains, hay, pasture and forest. Existing approaches for sampling optimization must be tested in this region and results from the different systems compared.
6. While this review primarily focused on determining soil C baseline levels, the next major challenge that should be addressed is how to optimize sampling for detecting soil C change. Future work should focus on understanding how variograms for soil C change relate to variograms for soil C status, and how land management and landscape

characteristics such soil type, slope, and manure application influence the spatial structure of soil C change.

7. Advancements in remote and proximal sensing techniques will change the way in which we approach sampling optimization, as both sampling and analytical costs will be tremendously reduced. Visable-near infrared (VNIR) and mid infrared (MIR) spectroscopy show great promise as quick, inexpensive, and nondestructive proxy measures for soil C (McCarty et al., 2002; Madari et al., 2005; Reeves III et al., 2006; Viscarra Rossel et al., 2006; Reeves III, 2010). Sampling costs can be greatly reduced with the use of tractor mounted "on-the-go" VNIR systems, which allow for spatially referenced, high-resolution measurements over large areas in a short amount of time. As accurate high-resolution soil C data become more available we will need to rethink how to integrate sampling for traditional analytical methods with these new sources of auxiliary data.

REFERENCES

Adams, W.A. 1973. The effect of organic matter on the bulk and true densities of some uncultivated podozolic soils. J. Soil Sci. 24(1): 10–17.

Allen, D.E., M.J. Pringle, K.L. Page, and R.C. Dalal. 2010. A review of sampling designs for the measurement of soil organic carbon in Australian grazing lands. Rangel. J. 32(2): 227–246.

Antle, J.M., S.M. Capalbo, S. Mooney, E.T. Elliott, and K.H. Paustian. 2002. A comparative examination of the efficiency of sequestering carbon in US agricultural soils. Am. J. Altern. Agric. 17(3): 109–115.

Antle, J., S. Capalbo, S. Mooney, E. Elliott, and K. Paustian. 2003. Spatial heterogeneity, contract design, and the efficiency of carbon sequestration policies for agriculture. J. Environ. Econ. Manage. 46(2): 231–250.

Bartholomeus, H.M., M.E. Schaepman, L. Kooistra, A. Stevens, W.B. Hoogmoed, and O.S.P. Spaargaren. 2008. Spectral reflectance based indices for soil organic carbon quantification. Geoderma 145(1–2): 28–36.

Bellamy, P.H., P.J. Loveland, R.I. Bradley, R.M. Lark, and G.J.D. Kirk. 2005. Carbon losses from all soils across England and Wales 1978-2003. Nature 437(7056): 245–248.

Bilgili, A.V., F. Akbas, and H.M. van Es. 2011. Combined use of hyperspectral VNIR reflectance spectroscopy and kriging to predict soil variables spatially. Precis. Agric. 12(3): 395–420.

Bishop, T.F.A., and A.B. McBratney. 2001. A comparison of prediction methods for the creation of field-extent soil property maps. Geoderma 103(1–2): 149–160.

Brus, D.J., and J.J. De Gruijter. 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). Geoderma 80(1): 1–44.

Brus, D.J., B. Kempen, and G.B.M. Heuvelink. 2011. Sampling for validation of digital soil maps. Eur. J. Soil Sci. 62(3): 394–407.

Burke, I.C., C.M. Yonker, W.J. Parton, C. V Cole, D.S. Schimel, and K. Flach. 1989. Texture, climate, and cultivation effects on soil organic matter content in US grassland soils. Soil Sci. Soc. Am. J. 53(3): 800–805.

Cambardella, C.A., T.B. Moorman, T.B. Parkin, D.L. Karlen, J.M. Novak, R.F. Turco, and A.E. Konopka. 1994. Field-scale variability of soil properties in central Iowa soils. Soil Sci. Soc. Am. J. 58(5): 1501–1511.

Cambule, A.H., D.G. Rossiter, and J.J. Stoorvogel. 2013. A methodology for digital soil mapping in poorly-accessible areas. Geoderma 192(0): 341–353.

Cambule, A.H., D.G. Rossiter, J.J. Stoorvogel, and E.M.A. Smaling. 2014. Soil organic carbon stocks in the Limpopo National Park, Mozambique: Amount, spatial distribution and uncertainty. Geoderma 213(0): 46–56.

Chatterjee, A., R. Lal, L. Wielopolski, M.Z. Martin, and M.H. Ebinger. 2009. Evaluation of different soil carbon determination methods. CRC. Crit. Rev. Plant Sci. 28: 164–178.

Chinembiri, T.S., M.C. Bronsveld, D.G. Rossiter, and T. Dube. 2013. The precision of C stock estimation in the Ludhikola watershed using model-based and design-based approaches. Nat. Resour. Res. 22(4): 297–309.

Conant, R.T., S.M. Ogle, E.A. Paul, and K. Paustian. 2011. Measuring and monitoring soil organic carbon stocks in agricultural lands for climate mitigation. Front. Ecol. Environ. 9: 169–173.

Conant, R.T., G.R. Smith, and K. Paustian. 2003. Spatial variability of soil carbon in forested and cultivated sites. J. Environ. Qual. 32(1): 278–286.

Cressie, N. 1990. The origins of kriging. Math. Geol. 22(3): 239–252.

Cressie, N.A.C. 1991. Statistics for Spatial Data. Wiley, New York.

Curtis, R.O., and B.W. Post. 1964. Estimating bulk density from organic-matter content in some Vermont forest soils. Soil Sci. Soc. Am. J. 28(2): 285–286.

Don, A., J. Schumacher, M. Scherer-Lorenzen, T. Scholten, and E.-D. Schulze. 2007. Spatial and vertical variation of soil carbon at two grassland sites—implications for measuring soil carbon stocks. Geoderma 141(3): 272–282.

Eswaran, H., E. Van Den Berg, and P. Reich. 1993. Organic carbon in soils of the world. Soil Sci. Soc. Am. J. 57(1): 192–194.

Gelder, B.K., R.P. Anex, T.C. Kaspar, T.J. Sauer, and D.L. Karlen. 2011. Estimating soil organic carbon in Central Iowa using aerial imagery and soil surveys. Soil Sci. Soc. Am. J. 75(5): 1821–1828.

Gregorich, E.G., C.F. Drury, and J.A. Baldock. 2001. Changes in soil carbon under long-term maize in monoculture and legume-based rotation. Can. J. Soil Sci. 81(1): 21–31.

De Gruijter, J.J. 2006. Sampling for natural resource monitoring. Springer, Berlin; New York.

Guo, L.B., and R.M. Gifford. 2002. Soil carbon stocks and land use change: a meta analysis. Glob. Chang. Biol. 8(4): 345–360.

Hengl, T., G.B.M. Heuvelink, and D.G. Rossiter. 2007. About regression-kriging: From equations to case studies. Comput. Geosci. 33(10): 1301–1315.

Hengl, T., G.B.M. Heuvelink, and A. Stein. 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120(1–2): 75–93.

Holmes, K.W., A. Wherrett, A. Keating, and D. V Murphy. 2012. Meeting bulk density sampling requirements efficiently to estimate soil carbon stocks. Soil Res. 49 (8): 680–695.

Hudson, B.D. 1994. Soil organic matter and available water capacity. J. Soil Water Conserv. 49(2): 189–194.

Jenkinson, D.S., S.P.S. Andrew, J.M. Lynch, M.J. Goss, and P.B. Tinker. 1990. The turnover of organic carbon and nitrogen in soil [and discussion]. Philos. Trans. R. Soc. London. Ser. B Biol. Sci. 329(1255): 361–368.

Kerry, R., and M. Oliver. 2003. Variograms of ancillary data to aid sampling for soil surveys. Precis. Agric. 4(3): 261–278.

Kerry, R., and M.A. Oliver. 2007. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. Geoderma 140(4): 383–396.

Knotters, M., D.J. Brus, and J.H. Oude Voshaar. 1995. A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. Geoderma 67(3–4): 227–246.

Kravchenko, A.N. 2003. Influence of spatial structure on accuracy of interpolation methods. Soil Sci. Soc. Am. J. 67(5): 1564–1571.

Kravchenko, A.N., and G.P. Robertson. 2007. Can topographical and yield data substantially improve total soil carbon mapping by regression kriging? Agron. J. 99(1): 12–17.

Kravchenko, A.N., G.P. Robertson, S.S. Snap, and A.J.M. Smucker. 2006. Using information about spatial variability to improve estimates of total soil carbon. Agron. J. 98(3): 823–829.

Lal, R. (Ed). Assessment Methods for Soil Carbon. Boca Raton, Fla. : Lewis Publishers.

Lal, R. 2004. Soil carbon sequestration impacts on global climate change and food security. Science (80-.). 304(5677): 1623–1627.

Lark, R.M. 2000. Estimating variograms of soil properties by the method-of-moments and maximum likelihood. Eur. J. Soil Sci. 51(4): 717–728.

Lark, R.M. 2009. Estimating the regional mean status and change of soil properties: two distinct objectives for soil survey. Eur. J. Soil Sci. 60(5): 748–756.

Madari, B.E., J.B. Reeves III, M.R. Coelho, P.L.O.A. Machado, H. De Polli, R.M. Coelho, V.M. Benites, L.F. Souza, and G.W. McCarty. 2005. Mid and near infrared spectroscopic determination of carbon in a diverse set of soils from the brazilian national soil collection. Spectrosc. Lett. 38(6): 721–740.

Maillard, É., and D.A. Angers. 2013. Animal manure application and soil organic carbon stocks: a meta-analysis. Glob. Chang. Biol.: n/a–n/a.

Matheron, G. 1962. Trait de geostatistique applique 1. 1. de. Technip, Paris.

McBratney, A.B., M. de L. Mendonça Santos, and B. Minasny. 2003. On digital soil mapping. Geoderma 117(1): 3–52.

McCarty, G.W., J.B. Reeves, V.B. Reeves, R.F. Follett, and J.M. Kimble. 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. Soil Sci. Soc. Am. J. 66(2): 640–646.

Minasny, B., A.B. McBratney, B.P. Malone, and I. Wheeler. 2013. Digital mapping of soil carbon. Adv. Agron. 118: 1–47.

Montes, F., M.J. Hernández, and I. Cañellas. 2005. A geostatistical approach to cork production sampling estimation in Quercus suber forests. Can. J. For. Res. 35(12): 2787–2796.

Mooney, S., K. Gerow, J. Antle, S. Capalbo, and K. Paustian. 2007. Reducing standard errors by incorporating spatial autocorrelation into a measurement scheme for soil carbon credits. Clim. Change 80(1-2): 55–72.

Nichols, J.D. 1984. Relation of organic carbon to soil properties and climate in the southern Great Plains. Soil Sci. Soc. Am. J. 48(6): 1382–1384.

Ogle, S.M., F.J. Breidt, and K. Paustian. 2005. Agricultural management impacts on soil organic carbon storage under moist and dry climatic conditions of temperate and tropical regions. Biogeochemistry 72(1): 87–121.

Pautsch, G.R., L.A. Kurkalova, B.A. Babcock, and C.L. Kling. 2001. The efficiency of sequestering carbon in agricultural soils. Contemp. Econ. Policy 19(2): 123–134.

Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30(7): 683–691.

Post, W.M., and K.C. Kwon. 2000. Soil carbon sequestration and land-use change: processes and potential. Glob. Chang. Biol. 6(3): 317–327.

Pringle, M.J., D.E. Allen, R.C. Dalal, J.E. Payne, D.G. Mayer, P. O'Reagain, and B.P. Marchant. 2011. Soil carbon stock in the tropical rangelands of Australia: Effects of soil type and

grazing pressure, and determination of sampling requirement. Geoderma 167–168(0): 261–273.

Puget, P., and R. Lal. 2005. Soil organic carbon and nitrogen in a Mollisol in central Ohio as affected by tillage and land use. Soil Tillage Res. 80(1): 201–213.

Reeves III, J.B. 2010. Near-versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done? Geoderma 158(1): 3–14.

Reeves III, J.B., R.F. Follett, G.W. McCarty, and J.M. Kimble. 2006. Can near or mid-infrared diffuse reflectance spectroscopy be used to determine soil carbon pools? Commun. Soil Sci. Plant Anal. 37(15-20): 2307–2325.

Sherrod, L.A., G. Dunn, G.A. Peterson, and R.L. Kolberg. 2002. Inorganic carbon analysis by modified pressure-calcimeter method. Soil Sci. Soc. Am. J. 66(1): 299–305.

Simbahan, G.C., and A. Dobermann. 2006. Sampling optimization based on secondary information and its utilization in soil carbon mapping. Geoderma 133(3): 345–362.

Simbahan, G.C., A. Dobermann, P. Goovaerts, J. Ping, and M.L. Haddix. 2006. Fine-resolution mapping of soil organic carbon based on multivariate secondary data. Geoderma 132(3): 471–489.

Singh, K., B.W. Murphy, and B.P. Marchant. 2013. Towards cost-effective estimation of soil carbon stocks at the field scale. Soil Res. 50 (8): 672–684.

Sommerfeldt, T.G., C. Chang, and T. Entz. 1988. Long-term annual manure applications increase soil organic matter and nitrogen, and decrease carbon to nitrogen ratio. Soil Sci. Soc. Am. J. 52(6): 1668–1672.

Syswerda, S.P., A.T. Corbin, D.L. Mokma, A.N. Kravchenko, and G.P. Robertson. 2011. Agricultural management and soil carbon storage in surface vs. deep layers. Soil Sci. Soc. Am. J. 75(1): 92–101.

Team, R.D.C. 2005. R: A language and environment for statistical computing. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2013. url: http://www. R-project. org.

Viscarra Rossel, R.A., D.J.J. Walvoort, A.B. McBratney, L.J. Janik, and J.O. Skjemstad. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131(1): 59–75.

Webster, R., and M.A. Oliver. 1992. Sample adequately to estimate variograms of soil properties. J. Soil Sci. 43(1): 177–192.

Webster, R Oliver, M. A.,, R. 2001. Geostatistics for Environmental Scientists. John Wiley & Sons, Chichester, England; New York.

West, T.O., and W.M. Post. 2002. Soil organic carbon sequestration rates by tillage and crop rotation. Soil Sci. Soc. Am. J. 66(6): 1930–1946.

West, L.T., S.W. Waltman, S. Wills, T.G. Reinsch, E.C. Benham, C.S. Smith, and R. Ferguson. 2010. Soil carbon stocks in the US: Current data and future inventories. p. 14. *In* Proceedings of International Workshop on Evaluation and Sustainable Management of Soil Carbon Sequestration in Asian Countries. Bogor, Indonesia.

Worsham, L., D. Markewitz, and N. Nibbelink. 2010. Incorporating spatial dependence into estimates of soil carbon contents under different land covers. Soil Sci. Soc. Am. J. 74(2): 635–646.

Xiao, X., G. Gertner, G. Wang, and A. Anderson. 2005. Optimal sampling scheme for estimation landscape mapping of vegetation cover. Landsc. Ecol. 20(4): 375–387.

Zhang, S., Y. Huang, C. Shen, H. Ye, and Y. Du. 2012. Spatial prediction of soil organic matter using terrain indices and categorical variables as auxiliary information. Geoderma 171–172(0): 35–43.

Chapter 2

Sampling Optimization for Soil Carbon Assessment in a Complex Agroecosystem of the Northeastern United States

Abstract

Low-cost accurate methods for estimating soil carbon (C) stocks are needed if terrestrial C offset markets are going to be implemented in the United States. Accurately measuring C stocks is often prohibitively expensive due to high spatial variability and analytical costs, therefore the development of cost–effective sampling designs and methods of inference are critical. We evaluated sampling optimization approaches for estimating soil C baseline levels for a dairy farm in Harford, NY with multiple land uses, including cultivation of silage corn, alfalfa hay, pasture, and forest. Three hundred and nineteen samples were collected in a spatially balanced design over a 232 hectare area to a depth of 30 cm. Secondary variables including soil type, elevation, slope, cropping history, and manure application rate were assessed for correlations with soil C and suitability for sampling stratification. Random, stratified, and systematic sampling arrangements at three sampling densities (n = 253, 160, 83), were compared to the full sampling grid (n = 319) using both design-based and model-based approaches for soil C assessment. For the model-based approach ordinary kriging (OK) was performed using variograms fit by both method of moments (MOM) and residual maximum likelihood (REML). Total soil C stocks for the sampling area were estimated by three different approaches: i) spatial mean (SM) where total C stocks are calculated by the area-weighted average of the mean C stocks for each landscape unit; ii) ordinary kriging where the sum of the predicted values for the interpolation grid are

used to determine total C-stocks; iii) (SSURGO) where average C stocks are based on estimates from the Soil Survey Geographic database with total C stocks calculated from the area-weighted average for each soil map unit. Nearly all secondary variables showed minimal correlations with soil C stocks. The systematic sampling arrangement was preferred over random or stratified arrangements because RMSE increased little with the reduced sample size, and the distribution of soil C stocks for the lowest sampling density closely resembled the full sampling grid. Landscape units defined by manure application rate explained the spatial variability of soil C-stocks better than any other categorical variable. Cropland categorized by high (98.1 Mg ha^{-1}) manure application rates had significantly greater mean C stocks than medium (88.3 Mg ha^{-1}) and low (87.9 Mg ha^{-1}) categories, while forest (91.6 Mg ha^{-1}) and pasture (91.3 Mg ha^{-1}) showed no significant differences. These trends imply that intensive manure application rates may be increasing soil C stocks to levels equal to or greater than the relatively undisturbed native vegetation in this region. Model-based approaches provided more reliable estimates for soil C stocks than design-based approaches. SM resulted in a higher RMSE than OK, 20.7 and 23.1 Mg ha^{-1} compared to 18.0 and 22.4 Mg ha^{-1}, respectively. Additionally, when the sampling density was reduced from 319 to 83, OK estimates fluctuated less than SM, with mean and total soil C stocks for the entire farm differing by 2% from that of the full sampling grid. Estimates of total C stocks to 30 cm for the entire 232 ha sampling area ranged from 16217–20049 Mg. Model-based approaches provided the most reliable estimates of soil C stocks. SSURGO based estimates consistently underestimated soil C stocks by 2.6-18.1 % compared to the full grid sampling, but given the low cost this approach may be of interest in some circumstances.

1. Introduction:

Soil has a tremendous potential for storing atmospheric C, with managed ecosystems having a storage capacity of 55 to 78 GT (Lal, 2004). It has been estimated that soil represents the world's largest stock of terrestrial C (1500Pg), nearly twice the amount of C that is stored in the atmosphere, fauna, and vegetation combined (Bartholomeus et al., 2008). The conversion of natural ecosystems to intensively managed agricultural lands has resulted in a global reduction of soil organic C (SOC) (Post and Kwon, 2000; West and Post, 2002; Ogle et al., 2005). Recent studies by the United States Environmental Protection Agency (EPA) have estimated that changes in management by the agricultural and forestry sectors have the capacity to reduce the annual U.S GHG emissions by 10-25 percent (EPA 2005).

Market based C trading has been proposed as a mechanism for incentivizing agricultural management practices that store C in the soil (Pautsch et al., 2001; Antle et al., 2002; Conant et al., 2011). Under a market based trading system farmers would be paid for the amount of soil C stored due to some change in soil management. Methods for determining offset payments can be based on actual sampling and soil C measurements, biophysical process models, or assigned rates of change for different management practices. While adoption of specific management practices is known to affect soil C, market based trading for soil C is likely to require that payments be based on the actual quantity of C stored (Antle et al., 2003). Soil C is typically extremely variable at both large and small spatial scales, thus accurately determining soil C levels over a large area can be prohibitively expensive due to high sampling and analytical costs (Conant et al., 2003).

Developing low cost approaches for evaluating SOC baselines and changes in SOC stock, from the field to the global scale has been identified as a high priority by researchers and policy-makers (Conant et al., 2011). Approaches for soil C assessment can be divided into two fundamentally distinct categories: design-based and model-based approaches (de Gruijter et al., 2006). Design-based approaches, often used for classical survey sampling, regard the population of values in a region as fixed while randomness is introduced through the selection of sampling locations. Furthermore, when calculating weighted averages, weights are based on selection probabilities rather than the geographic location of samples. In model-based approaches, values of a region are thought of as just one realization of a stochastic model, and weights are based on the covariance between observations as determined by the geographic location of the observations. In the model-based approach, sampling locations need not be random because randomness has been introduced through the model (de Gruijter et al., 2006).

Geostatistical model-based approaches rely on the concept of spatial autocorrelation, first introduced by Matheron (1962), to develop methods for predicting values at un-sampled locations in space. The underlying assumption is samples located closer in space will be more similar than those that are further away. The degree of spatial autocorrelation can be quantified by the experimental semi-variogram. Variograms are most commonly estimated using Matheron's (Matheron, 1962) methods of moments variogram (MoM) equation, but sample sizes of 100 to 150 are required to adequately estimate the variogram (Webster and Oliver, 1992). This relatively large initial sample requirement has led some soil scientists to stray away from model-based approaches to soil C assessment. Various authors have shown that the range of spatial autocorrelation can be used to determine the sampling interval so that redundant information is minimized (Kerry and Oliver, 2003). In cases where the range is large relative to the size of a

field or farm, the required sample size may not be large enough to accurately estimate the variogram. Computing variograms by a residual maximum likelihood (REML) approach has been proposed as a solution to the large sample requirement of MoM variograms. Kerry and Oliver, (2007) compared MoM and REML variograms for soil clay content at four field sites in England. It was found that predictions based on REML variograms were generally more accurate than those of MoM variograms when there were fewer than 100 sampling sites. They suggest that a sample size of around 50 samples with an appropriate spacing is adequate for kriging soil properties, for both purposes of precision agriculture and detection of soil contaminates.

In a recent review on sampling designs for measuring soil C in Australian grazing lands, Allen *et al*., (2010) makes the case that design-based sampling and inference is more efficient than geostatistical (i.e., model-based) approaches for measurement of soil C. Similarly, McKenzie *et al*. (2000) proposed a stratified random sampling approach with a minimum of 4 replicates per strata for soil C accounting in Australia at the national scale. Singh *et al*., (2013) developed a framework for quantifying sampling costs and attempted to determine the most efficient strategy for estimating soil C stocks at the field scale. Soil C displayed weak spatial correlation and correlation with auxiliary variables, thus stratification and model-based inference provided little benefit.

Mooney *et al*. (2007) recommended stratified random sampling and found that spatial autocorrelation could be used to reduce standard errors and narrow confidence intervals for sample estimates of the mean quantity of soil C sequestered per hectare within a region. This reduction in uncertainty could reduce transaction costs associated contracts for C-credits, which would in turn increase C offset payments to producers. A similar approach was taken by Worsham *et al*., (2010) where they evaluated how land cover affects the spatial structure of soil

C in forests in Georgia, USA, and how this will in turn influence prediction error and future sampling. They found that the standard error for mean soil C could be reduced when spatial autocorrelation was accounted for, and sample size could be reduced without a loss of accuracy. However, in contrast to Mooney et al. (2007), Worsham et al. (2010) found that a systematic sampling protocol was more efficient than using a randomized approach. Xiao *et al*, (2005) compared the efficiency of design-based and model-based approaches in terms of cost and accuracy for estimating the local and global mean vegetation cover in Fort Hood, Texas. They found that model-based approaches were more cost-efficient than classical designs because redundant information is reduced when spatial dependence is accounted for in the sample design.

While there is still debate on whether design-based or model-based approaches are most appropriate for soil C assessment, or whether stratified random or systematic sampling are most appropriate for developing semivariograms used in geostatistical modeling, it seems clear that incorporating the spatial structure of soil C in any analysis will improve results when sufficient spatial autocorrelation exists at the scale of the study site. To date, relatively few studies have focused specifically on sampling optimization for soil C assessment at the agroecosystem scale, and even fewer studies have evaluated how design-based inference compares to model-based inference with varying soil sampling optimization approaches (Brus and De Gruijter, 1997; Pringle et al., 2011).

The objectives of this study were to:

(i) compare the performance of design-based and model-based approaches for soil C assessment of a complex agroecostyem in the NE United States;

(ii) compare the performance of variograms estimated by REML and MOM for OK;

(iii) evaluate the effect of sample size and design (random, stratified, systematic) on variogram estimates and model performance for soil C assessment;

(iv) compare SSURGO estimated C stocks with design-based and model-based approaches.

2. Methods

2.1. Site description

The field site was at the Cornell University Animal Science Teaching and Research Center (T&R), located in Harford, NY (42.427° N, 76.228° W) (Figure. 1). The T&R is a 1052 ha working livestock farm with approximately 1000 dairy cattle, 500 beef cattle, and 500 meat and fiber sheep. The dairy unit of the farm is managed as a concentrated animal feeding operation where cattle are primarily housed in barns, while animals in the sheep and beef units are mostly put out to pasture. Major land uses on the farm include cultivation of silage maize, alfalfa hay, beef pasture, sheep pasture, and forest, with all cropped and hay fields receiving manure amendments. Crop land is mostly situated in the valley, while pasture is generally located in the surrounding hillsides. Soils at this site are defined by well drained medium textured glacial outwash deposits. Mean annual precipitation at this location is 956 mm, with surrounding native vegetation primarily consisting of mixed temperate deciduous and coniferous forest.

A 232 ha sampling area was randomly chosen within the farm, and it encompassed the majority of land use types, soil types, and topographical features on the farm. Crop land, beef pasture, sheep pasture, and forest accounted for 141 ha, 22 ha, 20 ha, and 15 ha, respectively. Soils belonged to the Alfisol order, with the dominant soil series in the sampling area consisting of Howard (139 ha, Loamy-skeletal, mixed, active, mesic Glossic Hapludalf), Langford (26 ha,

Fine-loamy, mixed, active, mesic Typic Fragiudepts), and Valois (13 ha, Coarse-loamy, mixed, superactive, mesic Typic Dystrudepts). Elevation of the sampling area ranged from 363 m to 441 m with a mean elevation of 381 m.

2.2. Sample design

Within the 232 ha sampling area, 319 core samples to a depth of 30 cm were collected during July 2012. Samples points were laid out in a systematic equilateral triangular grid with a distance of 103 m between points (n=240). A sample density of 103 m between points was chosen based on the minimum contiguous area that represented the major biophysical characteristics of the landscape (Figure 1). An equilateral triangle design has been shown to provide the most reliable estimate of semivariograms as the average distance to unsampled locations is minimized, and anisotropy is better detected (Yfantis et al. 1987; McBratney et al. 1981). For 1/4 of the points, an additional offset sample (n=79) was collected at a 1m distance from the original point in alternating directions (distance of 178.4 m between offset points). The 1 m offset samples provide valuable information on semivariance at short distances that cannot be estimated from the larger grid. Understanding short range variability is particularly important for estimating the variogram nugget, choosing variogram models, and geostatistical interpolation (De Gruijter, 2006). A Trimble Juno SD (Trimble Navigation, Ltd., Sunnyvale, CA, USA) global positioning system (GPS) was used to locate sampling points from a predefined sampling grid. When sample points fell on irrelevant land uses such as roads, buildings, or bodies of water, the sample was moved to the nearest relevant land use and its GPS coordinate was recorded.

2.3. Sampling method

Undisturbed soil cores to a depth of 30 cm were collected with either a JMC Environmental Sub Soil Probe Plus (Clements, Inc., Newtown, IA, USA) with a 2.81cm inner diameter steel probe driven by a slide hammer, or with steel probes built to similar specifications as the JMC and with the same cutting tip driven in to the soil by pounding the end of the probe with a wooden block and mallet. The degree of compaction was determined by measuring the distance between the top of the probe and the soil surface within the probe. Samples with more than 15 percent compaction were discarded, and a new sample was collected as close as possible. Upon 10 unsuccessful tries the sample point was omitted from the design. Twenty samples were removed due to sampling difficulties, but this caused little distortion to the overall grid. Debris such as leaf litter or crop residues was removed from the soil surface prior to sampling. Upon collection samples were immediately stored in an ice chest; at the end of the day all samples were brought back to the laboratory where they were stored at 4˚C until they could be processed for analysis. Samples were analyzed within 60 days.

2.4. Auxiliary variables:

Auxiliary data is often readily available or inexpensive to measure, and may provide valuable information about soil C variability which may be used to improve soil C predictions at un-sampled locations. Five auxiliary variables related to soil forming processes were acquired for the entire study area, and were examined for relationships with soil C.

2.4.1. Relative elevation

A 1/3 Arc Second 10 m resolution digital elevation model (DEM) was obtained from the U.S. Geological Survey (USGS), EROS Data Center. The DEM was then used to determine elevation and slope for the region (Fig. 1).

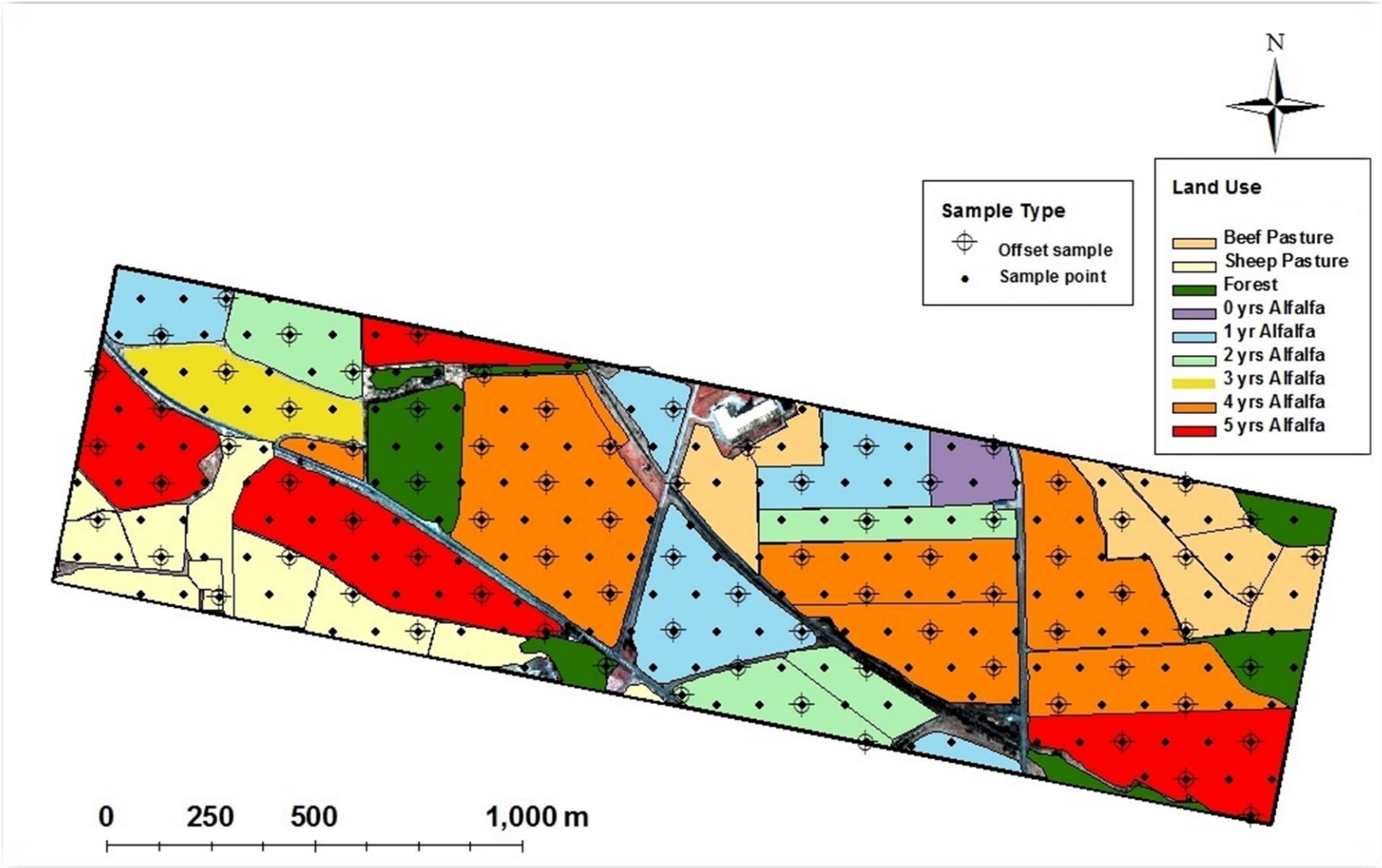


Fig. 1.a. Land use map for the 232ha sampling area, land uses included beef pasture, sheep pasture, forest, and crop land in a corn-alfalfa rotation with varying years of alfalfa. Both regular and offset samples are shown.

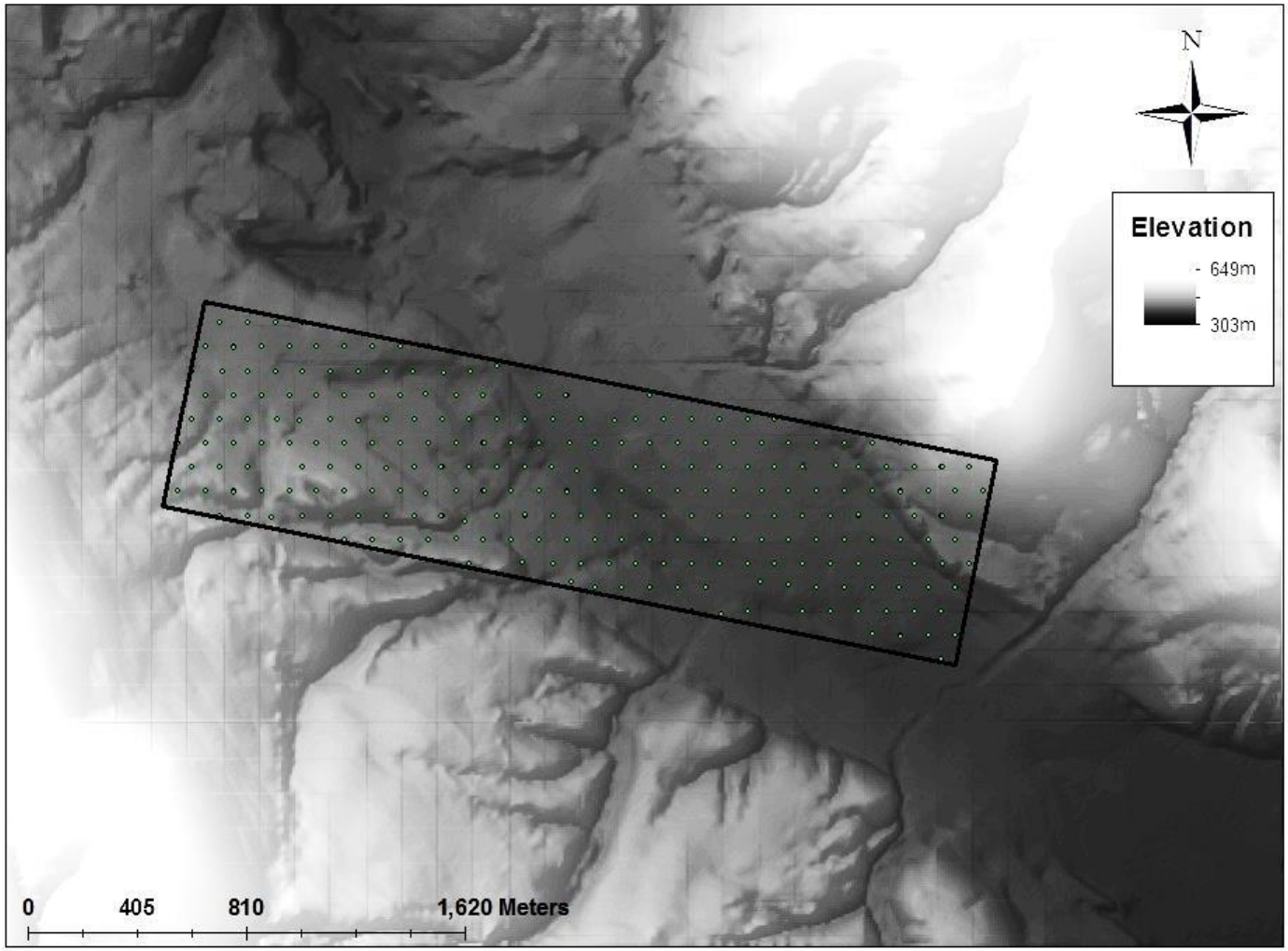


Fig. 1.b. Digital elevation map for the sampling area and the surrounding region. Sample locations are denoted by points on the map.

2.4.2. Permanent land use and crop rotation

Cropping history and land use records were compiled for the 2006 through 2011 growing seasons, and a digital land use map was created using ArcGIS 10.0 (ESRI Inc., Redlands, CA, USA). Land uses across the farm were categorized based on permanent land use type and crop rotation (Table 2, Table 6). Permanent land use categories included forest, pasture, and cropland. Crop rotations primarily consisted of maize-alfalfa rotations, with small grains occasionally incorporated in the rotation. As such crop rotation categories were classified by the number of years in alfalfa during the 2006 through 2011 growing seasons, where categories included forest, pasture, and years of alfalfa ranging from 0 to 5 years (Table 6).

4.3. Soil series

A digital soil map (1:12000) was downloaded from the national Soil Survey Geographic (SSURGO) database (USDA-NRCS) and was used to obtain soil series data for the entire sampling area. In addition to soil series data, attribute data was obtained for each soil map unit, including soil organic matter (SOM) and bulk density to a depth of 30 cm. Two soil samples were classified as a mucky silt loam and were removed from this analysis due to insufficient sample size.

2.4.4. Manure application

Manure application records from 2009-2012 were obtained for the sampling area (Table 1). Records consisted of daily application rate, type of manure, location, and total kjeldahl N content per application. Manure was analyzed for nutrient content and percent solids on an annual basis. Types of manure being applied included beef and dairy solids, beef and dairy

liquids, sheep solids, and compost. The various forms of manure were reclassified into three groups based on the total amount of liquid manure, solid manure, and total N applied (Table 1).The total amount of N applied to each field from 2009-2012 (TN) was used as a measure of manure application. Manure application rate was grouped into land use categories of high, medium, low, pasture, and forest. Forest and pasture received no manure, while high, medium, and low received 980-2456 kg N ha^{-1}, 344-632 kg N ha^{-1}, and 63-200 kg N ha^{-1}, respectively.

Table 1. Manure application categories as determined by the amount of liquid, solid, and total N (TN) manure applied to each field.

	Field ID	Area (ha)	Liquid (kl / ha)	Solid (Mg / ha)	TN (kg/ha)
High	10C	11.4	197	378	2456
High	6	5	133	299	1629
High	17	3.4	164	263	1556
High	20	4.6	256	116	1025
High	7	9.1	162	141	980
Med	14	5.4	0	105	632
Med	5A	3.2	377	4	581
Med	10B	8.4	212	47	580
Med	10A	11.7	159	65	556
Med	19B	5.9	112	52	468
Med	16	21.9	154	33	458
Med	27	7.8	73	56	410
Med	52	2.8	0	80	374
Med	8	8.7	117	51	369
Med	19C	1.1	173	0	344
Low	5B	7.2	97	14	200
Low	13	1.3	69	11	191
Low	19A	7.4	49	14	163
Low	31	13.4	60	5	140
Low	15	10	32	4	63

2.4.5. Correlations between auxiliary variables and soil C

Continuous auxiliary variables were evaluated by a Pearson's correlation matrix for relationships with soil C stocks (Table 4). Continuous variables included C concentration, bulk density, C stocks, elevation, slope, liquid manure, solid manure, and manure TN. Relationships between categorical variables and soil C stocks were examined by *post hoc* ANOVA where pairwise comparisons using a Student's t test were applied to each categorical grouping (Table 2,5,6,7). Categorical variables included permanent land use type, cropping history, soil texture class, and manure land use categories.

2.5. Soil analyses

Soil samples were air dried until friable at which point they were sieved to 2 mm and oven dried at 35˚C for a minimum of 48 hours. Prior to air drying a 20 g field moist subsample was removed and dried a 35˚C for 48 hrs, weighed and then dried at 105˚C for 48 hours to determine soil moisture content. Soil organic C (SOC) concentration was determined by measuring inorganic C (IOC) by pressure calcimeter method (Sherrod et al., 2002) and subtracting this from total C (TC). IOC was not detected in any of our samples, all mention of soil C is referring to soil organic C. TC was measured by dry combustion (VarioMax, Elementar). Bulk density for the fine-earth fraction was calculated as described by (Holmes et al., 2012) using the following formula:

$$BD = \frac{\text{mass}_{soil} - \text{mass}_{rock}}{\text{volume}_{soil} - \frac{\text{mass}_{rock}}{\text{density}_{rock}}} \quad (1)$$

Where mass_{soil} is the mass of the <2 mm fraction, volume_{soil} is the volume of the <2 mm fraction, mass_{rock} is the sample mass of the >2 mm fraction, and density_{rock} is assumed to be 2.6 g cm^{-3}. Soil moisture content was used to adjust mass_{soil} to a weight equivalent to drying at 105˚C.

Soil C stocks (C per unit area) to a depth of 30 cm were calculated with the following equation:

$$C\ stocks = BD * C_{conc} * Depth * (1 - CF) \quad (2)$$

Where *BD* is the bulk density of the fine-earth fraction (Eqn 1), C_{conc} is the C concentration, *Depth* is sampling depth, and *CF* is the fraction of coarse material >2 mm.

2.6. Sampling designs

Model performance and estimates of soil C were compared for different sample designs and sampling densities. Random, stratified random, and systematic sample designs for different sample sizes were created by removing 21, 50, and 74 % of samples from the initial sampling grid. Sample sizes for the 0, 21, 50, and 74% removed were 319, 235, 160, and 83, respectively. For the random design all samples were removed at random (Fig. 2). For the stratified random design stratification was based on manure land use categories, and the sample size for each land use was proportional to its respective area (Fig. 3). Samples within each land use were removed at random. For the systematic design both regular and offset samples were systematically removed such that the proportion of regular to offset samples remained constant for all sample sizes (Fig. 4).

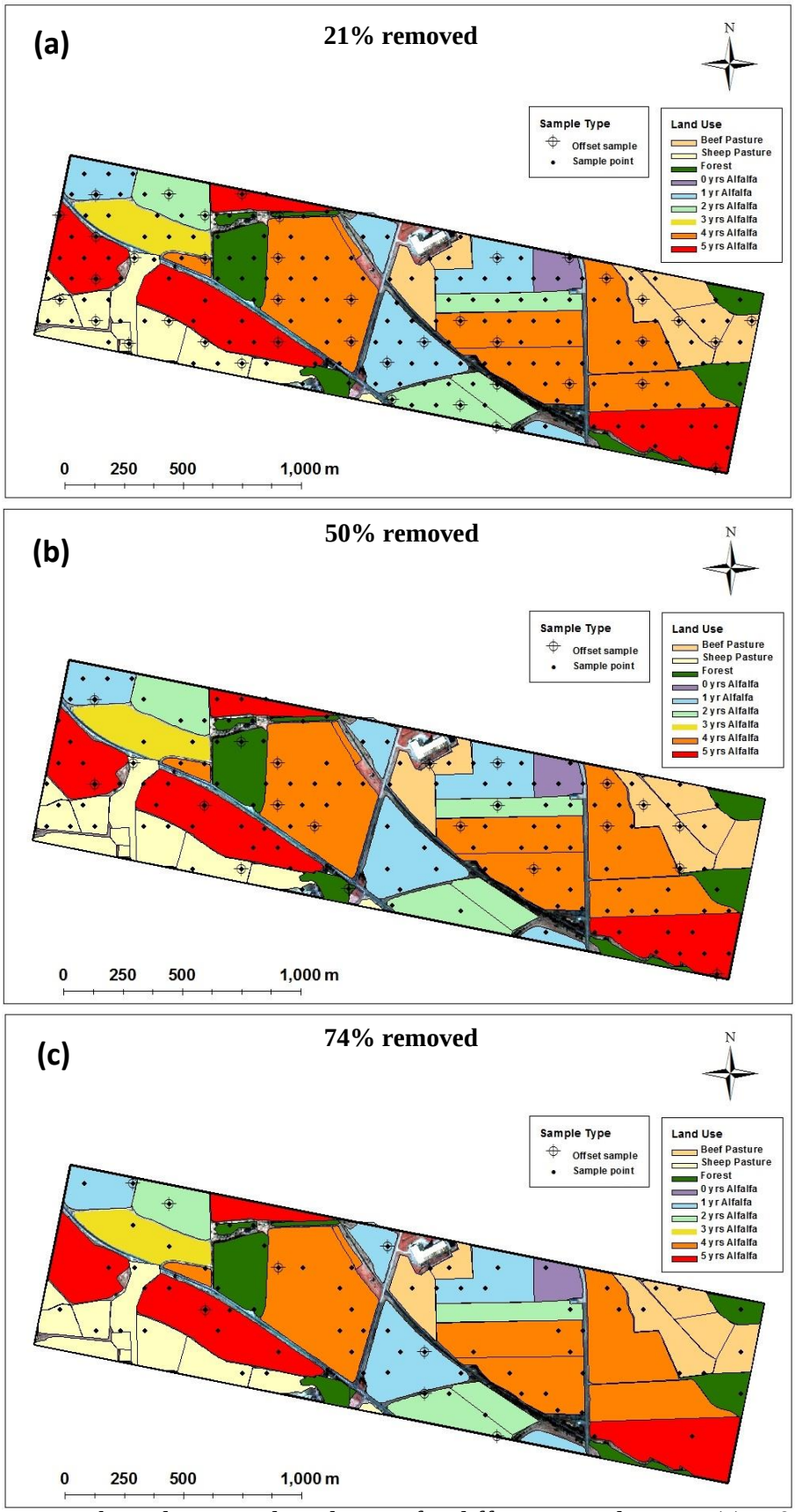


Fig.2. Land use maps with random sampling designs for different sample sizes. (a) 21% of samples removed (n=253), (b) 50% of samples removed (n=160), (c) 74% of samples removed (n=83)

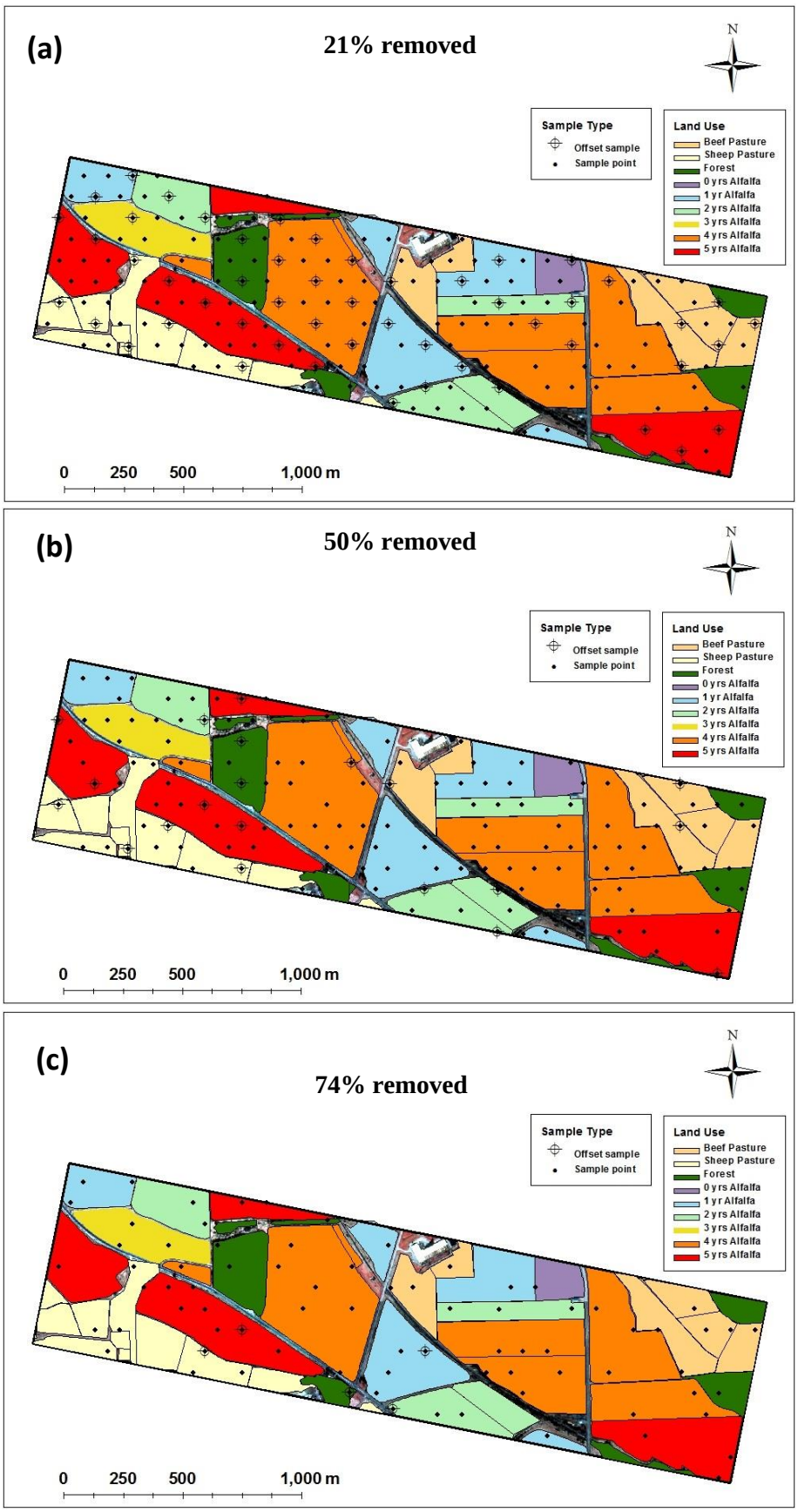


Fig.3. Land use maps with stratified random sampling designs for different sample sizes. (a) 21% of samples removed (n=253), (b) 50% of samples removed (n=160), (c) 74% of samples removed (n=83)

Systematic Sample Designs

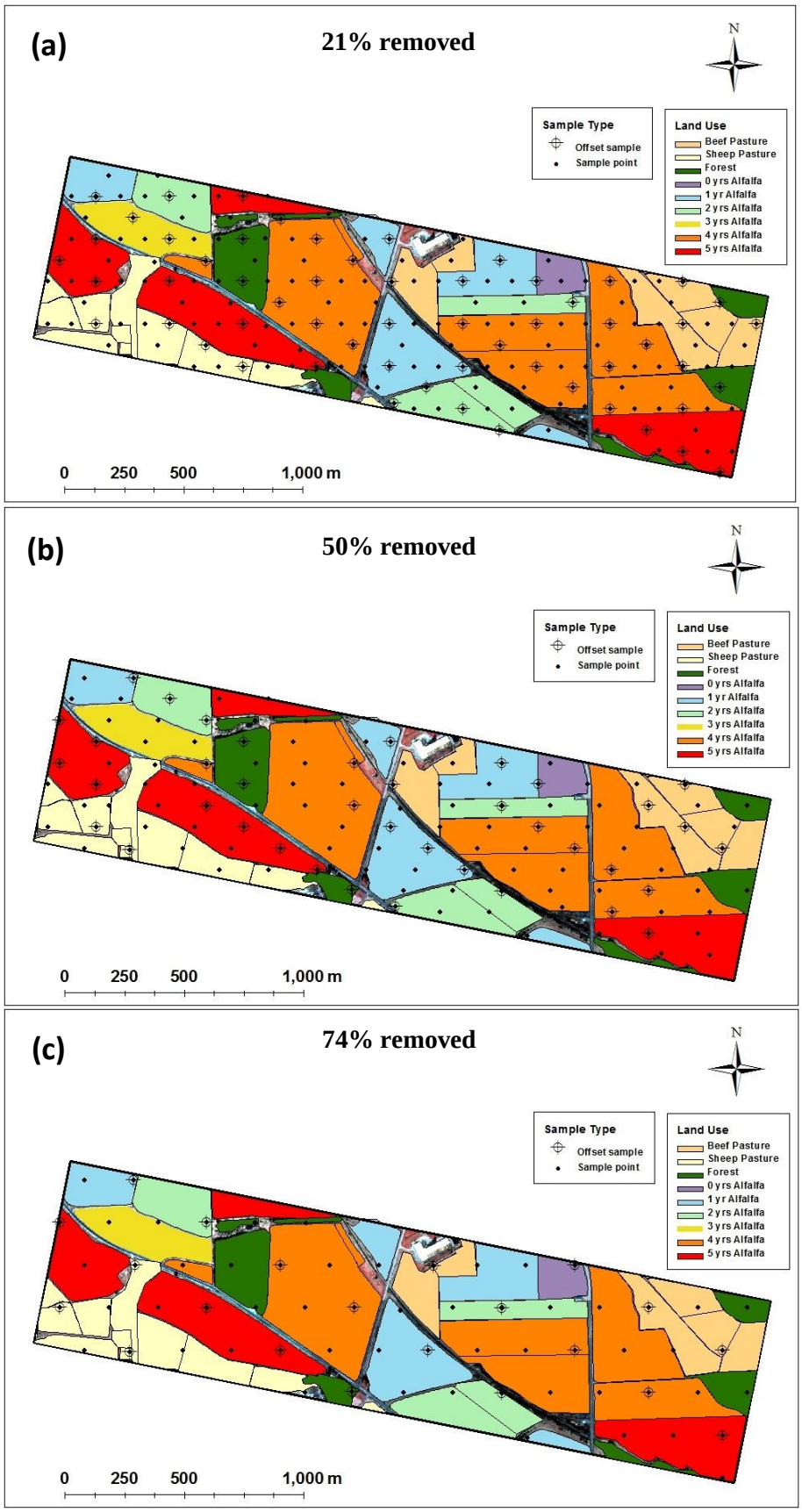


Fig.4. Land use maps with systematic sampling designs for different sample sizes. (a) 21% of samples removed (n=253), (b) 50% of samples removed (n=160), (c) 74% of samples removed (n=83)

2.7. Design-based inference

All statistical analyses for design-based inference were performed with JMP 10 (SAS Institute, Cary, NC, 2012). Under the design-based approach soil C concentration and stocks were estimated by the sample mean, sample variance, and variance of the sample mean [μ_s, σ^2_s, $\sigma^2_s(\mu_s)$] for all sampling arrangements and sampling densities:

$$\mu_s = \frac{1}{n}\sum_{i=1}^{n} y_i \quad (3)$$

$$\sigma_s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \mu_s)^2 \quad (4)$$

$$\sigma_s^2(\mu_s) = \frac{\sigma_s^2}{n} \quad (5)$$

Equations (1-3) were applied to the random and systematic sampling designs. For the stratified random sampling design the above equations were adjusted to account for the change in section probabilities resulting from stratification (Allen et al., 2010):

$$\mu_{st} = \sum_{h=1}^{H} a_h \mu_h \quad (6)$$

$$\sigma_{st}^2 = \left(\frac{1}{n}\sum_{i=1}^{n} y_i^2\right) - \mu_{st}^2 + \sigma_{st}^2(\mu_{st}) \quad (7)$$

$$\sigma_{st}^2(\mu_{st}) = \sum_{h=1}^{H} a_h^2 \sigma_h^2(\mu_h) \quad (8)$$

Where H is the number of strata; a_h is the proportion of the area representing the hth stratum; and μ_h is its mean. The square root of the sample variance is equivalent to the RMSE, and so this value was used to compare design-based and model-based approaches.

2.8. Model-based inference

All statistical analyses for model-based inference were performed in R 2.15.2 (R CoreTeam, 2005). Geostatistical analyses were carried out using the gstat R package (Pebesma, 2004). Semivariograms and OK predictions were determined for all sampling arrangements and sampling densities.

2.8.1. Spatial autocorrelation

Spatial autocorrelation describes the covariance between observations relative to their location in space. The degree of spatial autocorrelation can be quantified by the experimental semi-variogram (Figure. 6):

$$\bar{\gamma}(\mathbf{h}) = \frac{1}{2n(\mathbf{h})}\sum_{i=1}^{n(\mathbf{h})}\{z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})\}^2 \qquad (9)$$

Where $\bar{\gamma}(\mathbf{h})$ is the average semivariance as a function of **h**; $n(\mathbf{h})$ is the number of point pairs as a function of **h**; $z(\mathbf{x}_i)$ is the observed value of z at the ith location; and $z(\mathbf{x}_i + \mathbf{h})$ is the observed value of z at a distance of **h** from $z(\mathbf{x}_i)$. Variogram models were computed by method of moments and residual maximum likelihood.

The variogram is defined by three parameters: the range, the sill, and the nugget. The range can be thought of as the zone of influence, or the distance at which samples are no longer

spatially correlated. The sill represents the total observed variance, and the nugget represents the portion of the variance that cannot be explained by spatial correlation. The nugget is often the result of measurement error and spatial trends that occur at a scale smaller than the minimal sampling distance (Cressie, 1991).

2.8.2. Ordinary kriging

The purpose of kriging is to estimate the value of a random variable at unsampled locations, where predictions consist of the weighted average of nearby samples, and weights are assigned based on patterns in spatial variability as determined by the semivariogram model. Ordinary kriging is one of the simplest and most commonly used forms of kriging, and is given by the formula (Hengl et al., 2007):

$$\hat{z}(s_0) = \sum_{i=1}^{n} \lambda_i \cdot z(s_i) \qquad (10)$$

Where $\hat{z}(s_0)$ is the OK predicted value z at location (s_0), n is the number of samples in the search neighborhood, and λ_i is the weight assigned to the ith observation $z(s_i)$. Kriging is an optimal prediction method in that weights are assigned to sample points within a search neighborhood in such a way that the estimation or kriging variance $E[\{\hat{z}(s_0) - z(s_0)^2\}]$, is minimized, and that estimates are unbiased (Webster and Oliver., 2001).

2.8.4. Model evaluation

For each sampling design, density, and interpolation method the root mean squared error (RMSE) was used to evaluate model performance:

$$\text{RMSE} = \sqrt{\frac{1}{l} \cdot \sum_{j=1}^{l} [\hat{z}(s_j) - z^*(s_j)]^2} \qquad (11)$$

RMSE is calculated by comparing the predicted value $\hat{z}(s_j)$ at a validation sample location with an actual observation $z^*(s_j)$, where l is the number of validation samples. RMSE was determined by leave-one-out-cross-validation (LOOC). In the LOOC approach the model is calibrated with one sample removed, this sample will then be used to validate the model. The removed sample is then returned to calibration set, and the entire procedure is repeated until all samples have been removed.

2.9. Estimating total C stocks

Total C stocks for the entire sampling area were estimated using three different approaches (see Appendix A):

(1) spatial mean (SM): in this approach mean C stocks were estimated by design-based methods for each manure category and the entire farm. Mean C stocks Mg ha^{-1} for each category were then multiplied by its respective area to determine the total C stocks in Mg for each manure land use category.

(2) ordinary kriging (OK): for this approach the spatial structure of the observations were used to predict soil C stocks at 10 m resolution. The grid cells contained within each manure category were then summed to determine total C stocks.

(3) SSURGO estimated (SSURGO): for this approach C concentration to a depth of 30 cm was estimated for each soil map within the sampling area by multiplying the SSURGO estimated OM by 0.58, a commonly used conversion factor (Chatterjee et al., 2009). C stocks were then determined from the SSURGO estimated bulk density and C concentration. Total C stocks for each manure land use category was estimated by multiplying the area for each map unit

contained within a manure category by its respective C stock, and C stocks for each map unit were then summed.

2.10. Comparing design-based and model-based approaches

Design-based and model-based performance was evaluated based on the RMSE for estimates of mean C stocks for the different manure land use categories (Chinembiri et al., 2013; Cambule et al., 2014). For the different approaches the RMSE is calculated slightly differently, however, we believe this is an appropriate method for comparing design- and model-based approaches. The main difference is that for the design-based approach the sum of squares is based on the difference between an observation and the mean, while for the model-based approach the sum of squares is based on the difference between a model prediction and observation that has been removed by LOOC.

3. Results and Discussion

3.1. Soil C summary statistics for the entire farm

Table 2. Mean and coefficient of variation (CV) for C concentration (C%), bulk density (BD), and C stocks for each permanent land use category and the entire farm. Mean values in each column with the same letter are not significantly different, α=0.05

		C %		BD (g cm^{-3})		C stock (Mg $ha^{-1)}$	
	N	mean	CV	mean	CV	mean	CV
Forest	20	3.4a	33	1.03a	20	91.6a	27
Pasture	63	3.2a	30	1.11b	15	91.3a	27
Crop	236	2.9b	27	1.22b	12	90.3a	21
Entire farm	319	3.0	29	1.19	14	90.6	23

Soil C concentration for the entire farm ranged from 0.98 to 6.75% with a mean value of 3.00 %, bulk density ranged from 0.73 to 1.65 g m^{-3} with a mean value of 1.19 g m^{-3}, and C stocks ranged from 30 to 191 Mg ha^{-1} with a mean value of 91 Mg ha^{-1} (Table 2). Bulk density, C concentration, and C stocks were evaluated for normality and it was determined transformations were not necessary. C concentration was more variable than either C stocks or bulk density, with coefficients of variation (CV) equal to 29, 23, and 14%, respectively (Table 2). Percent C was far more variable than bulk density. Reducing the number of bulk density samples will have little effect on the power of C stock estimates, but will significantly reduce sampling costs.

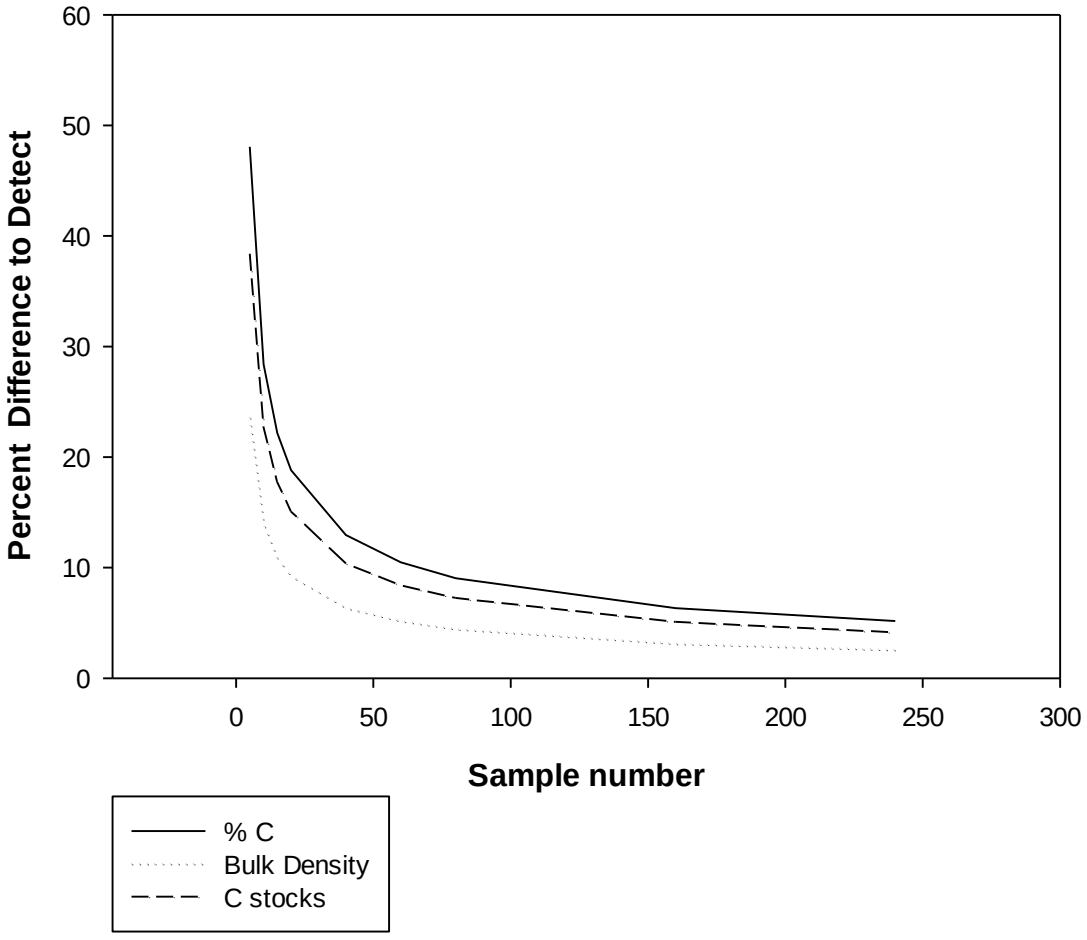


Fig. 5**.** Percent difference to detect for C concentration (C%), bulk density, and C stocks with increasing sample size for the entire farm, α=0.05.

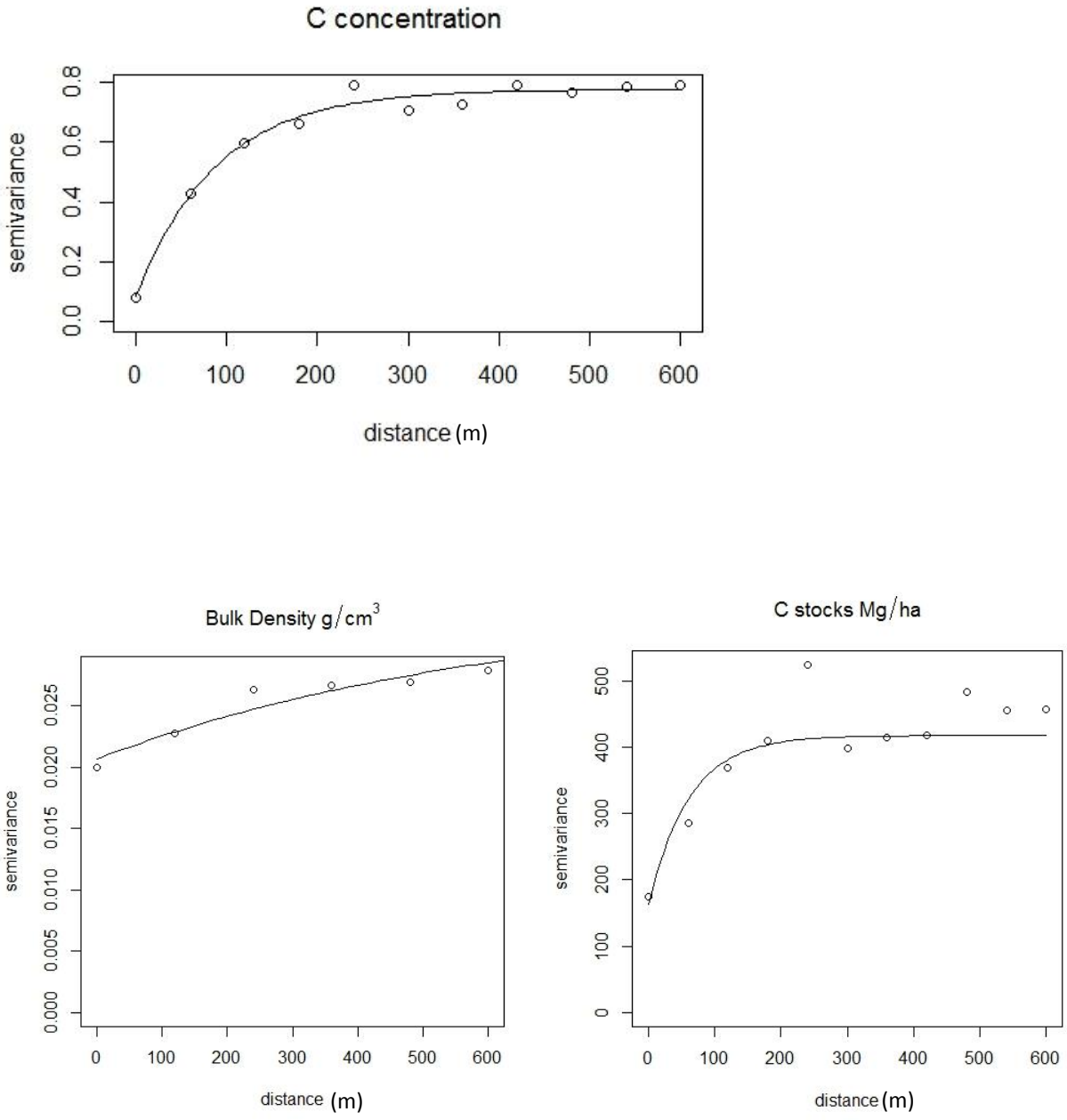


Fig.6. Experimental and fitted variograms for C concentration, bulk density, and C stocks for the entire farm n=319.

The strength of spatial autocorrelation varied considerably among the different soil properties (Fig. 6). C concentration showed the strongest spatial auto correlation with range of 206 m, and a nugget to sill ratio of 0.10. Bulk density displayed weak spatial auto correlation with a range of 410 m and a nugget to sill ratio of 0.74. C stocks displayed moderate spatial structure with a range of 226 m and a nugget to sill ratio of 0.40 (Table 3).

Table 3. Variogram model parameters sill, nugget, range, and nugget to sill ratio for C concentration (%C), bulk density (BD), and C-stocks.

	Sill	Nugget	Range (m)	Nugget : sill
C%	0.7247	0.08	206	0.10
BD	0.0267	0.02	410	0.75
C-stocks	429.7	171.91	226	0.40

The nugget to sill ratio provides a measure of the degree of spatial auto correlation, a ratio <0.25, 0.25-0.75, and >0.75 are categorized as having strong, moderate, and weak spatial structure, respectively (Cambardella et al., 1994). For both C concentration and C stocks it appears that the fitted variogram model is underestimating the sill and range of the experimental variogram, and so it is possible that the true range of spatial autocorrelation is on the order of 600 m (Fig. 6).

3.2. Evaluation of continuous auxiliary variables

All continuous auxiliary variables were evaluated for correlations with soil C stocks. Correlation coefficients between soil properties, topographical features and manure application rates are shown in Table 4. C concentration was negatively correlated with bulk density ($r = -0.53$, $P < .0001$). Building soil organic is a common strategy for reducing bulk density in agricultural systems, as such the relationship between soil C and bulk density is well established (Curtis and Post, 1964; Adams, 1973; Hudson, 1994). Correlations between C concentration and C stocks were significant ($r = 0.83$, $P<.0001$), while bulk density was not correlated with C

stocks. Of the explanatory variables, the amount solid manure applied showed the strongest relationship with soil C concentration and bulk density (r = 0.30, P<.0001) and (r = -0.25, P<.0001), respectively. From 2009-2012 the amount of solid manure applied to different fields across the farm ranged from 0 to 378 Mg ha^{-1}, time of application and farm management practices will greatly influence the retention on soil C from manure, but such large additions of biomass will inevitably raise soil C levels.

Table 4. Pearson's correlation coefficients (*r*) between soil properties, topographical features, and manure application rates. Variables included C concentration (%C), bulk density (BD), C stocks, elevation (Elv), slope, liquid manure application rate (Liquid), solid manure application rate (Solid), and amount of total nitrogen applied as manure (TN). Values marked by (*) are significant at $P \leq 0.01$ significance level.

	%C	BD ($g\ cm^3$)	C stocks ($Mg\ ha^{-1}$)	Elv (m)	Slope (%)	Liquid (kl)	Solid (Mg)	TN (kg)
% C	1.00	-0.53*	0.83*	-0.18*	-0.11	0.05	0.30*	0.27*
BD		1.00	-0.08	-0.04	-0.11	-0.12	-0.25*	-0.26*
C stocks			1.00	-0.16	-0.18*	0.00	0.17*	0.14
Elv				1.00	0.61*	-0.24*	-0.39*	-0.41*
Slope					1.00	-0.29*	-0.30*	-0.32*
Liquid						1.00	0.25*	0.43*
Solid							1.00	0.97*
TN								1.00

Terrain attributes derived from DEMs are some of the most commonly used covariates for prediction of soil C, as slope and elevation influence erosion and deposition processes (Minasny et al., 2013). Across the farm, slope and elevation represented a relatively large range of values from 0 to39% and 363 to 441m. However, both variables were positively skewed which may explain the observed low correlations with soil C. While it is likely that slope, elevation, and manure application rates are influencing soil C stocks through plant growth and organic matter inputs, correlations between auxiliary variables and soil C stocks are weak,

meaning these additional variables will likely provide little predictive power to linear interpolation models.

Both slope and elevation were negatively correlated with manure application. This is to be expected as crop land on this farm is generally located in flat areas with lower elevation. Additionally, field accessibility, particularly during the winter when much of the manure is applied, dictates where a farmer will apply manure on a given day. Fields with greater slope and elevation may be more susceptible to nutrient runoff and may be less accessible to manure spreading equipment when the ground is wet or snow-covered. For correlations with manure variables (Liquid, Solid, TN), pasture and forest land uses were removed from the analyses, as these land uses received no manure applications, but are expected to have a higher soil C content which would reduce the effect of a positive correlation with manure application rate (Guo and Gifford, 2002; Maillard and Angers, 2013).

3.3. Evaluation of categorical auxiliary variables

Permanent land use, cropping history, soil texture class, and manure application were evaluated for relationships with soil C stocks by *post hoc* ANOVA (Tables 2,5,6,7).

3.3.2. Permanent land use type

Land uses across the farm were grouped into three broad categories—crop land, pasture, and forest (Table 2). These groupings represent distinct management practices which do not change from year to year, and have been shown to influence soil C concentration and C stocks

(Post and Kwon, 2000; Gregorich et al., 2001; Ogle et al., 2005). Forest and pasture were more variable than cropland, showing higher CVs for C concentration, bulk density, and C stocks. Forest (3.4%) and pasture (3.2%) had significantly higher C concentration than cropland (2.9%). Forest had a significantly lower bulk density (1.03 g m^{-3}) than either pasture (1.11 g m^{-3}) or crop land (1.22 g m^{-3}). C stocks ranged from 90.3 to 91.6 Mg ha^{-1}, and there were no significant differences between permanent land use types.

Soil C stocks can be highly variable from site to site, levels at the T&R are comparable to many previous studies, but unlike those studies permanent land use type does not explain the spatial distribution of C stocks across the farm (Eswaran et al., 1993; Post and Kwon, 2000; West and Post, 2002; Ogle et al., 2005; Puget and Lal, 2005; Don et al., 2007). Puget and Lal (2005), found that forest (127 Mg ha^{-1}) had higher C stocks than pasture (107 Mg ha^{-1}) or a corn-soybean rotation (88 Mg ha^{-1}) in the top 30 cm for a Mollisol soil in Ohio. In a study comparing 30 important agricultural soils in New Jersey, it was shown that cropland on average had 49% lower C stocks than forest sites on the same soil type (Lal, 2001). It has been widely documented that changes in land use can have a strong effect on soil C stocks, with the general belief that intensive tillage from agriculture reduces soil C stocks as compared to native vegetation or pasture (Post and Kwon, 2000; Gregorich et al., 2001; West and Post, 2002; Ogle et al., 2005). However, most studies evaluating the effects of land use change on soil C stocks have not included agricultural systems with intensive manure application as was the case at our study site, which may explain the observed inconsistencies.

3.3.3. Cropping history

Cropping records from the 2006 to 2011 growing seasons were compiled and crop land was classified based on the number of years in alfalfa during the five growing seasons (Table 5). Alfalfa is a deep rooted perennial legume that has been shown to increase soil C stocks when compared to continuous maize (Gregorich et al., 2001). Categorizing crop rotation by years in alfalfa was an attempt to capture the non-uniformity of crop rotations between fields, and account for possible alfalfa effects. Maize silage and alfalfa are the predominant crops on this farm, but wheat was grown in field 27 during the 2009 growing season.. Forest (3.4 %) had significantly higher C concentration (3.4%) than 1 yr (2.9 %), 4 yrs (3.0 %), or 5yrs (3.0 %) of alfalfa. Forest (1.03 g cm^{-3}) had a significantly lower bulk density (1.03 g cm3) than any of the cropland categories (1.19 - 1.29 g cm^{-3}) or the beef pasture category (1.14 g cm^{-3}). C stocks ranged from 88.8 to 94.0 Mg ha^{-1}, but there were no significant differences between any of the land uses when classified by cropping history.

Table 5. Mean and coefficient of variation (CV) for C concentration (%C), C stocks, and bulk density (BD) for land uses categorized by cropping history. Mean values in each column with the same letter are not significantly different, α=0.05

		C %		BD (g cm^{-3})		C-stock (Mg ha^{-1})	
	N	Mean	CV	mean	CV	mean	CV
0 yrs Alfalfa	7	2.8ab	25	1.25abc	20	94.0a	8
1 yr Alfalfa	40	2.9b	20	1.25abc	10	93.2a	17
2 yrs Alfalfa	32	3.0ab	36	1.24abc	11	93.1a	30
3 yrs Alfalfa	10	2.8ab	20	1.29abc	9	93.7a	18
4 yrs Alfalfa	93	3.0b	30	1.19bc	13	87.9a	23
5 yrs Alfalfa	54	2.9b	22	1.21abc	11	88.8a	15
Beef Pasture	34	3.1ab	31	1.14cd	16	90.5a	31
Sheep Pasture	29	3.2ab	28	1.07de	13	92.2a	23
Forest	20	3.4a	33	1.03e	20	91.6a	27
Entire farm	319	3.0	29	1.19	14	90.6	23

There were no apparent relationships between the number of years in alfalfa and soil C stocks. These findings are consistent with Syswerda *et al*, (2011) who found no differences in C stocks in the top 20 cm between a maize-soybean-wheat rotation (32 Mg ha^{-1}) and a perennial alfalfa crop (36 Mg ha^{-1}) during a 12 year replicated experiment. However, Gregorich *et al*, (2001) found that after a 35 year experiment continuous maize (92.1 Mg ha^{-1}) had lower C stocks than a legume-based rotation (103.7 Mg ha^{-1}) in the top 40 cm of soil. The effect of crop rotation on soil C stocks will vary with climatic conditions and soil type (Burke et al., 1989; Ogle et al., 2005). It is possible that at the T&R years of alfalfa grown was not an appropriate measure for categorizing cropping history, rotations were too similar to cause differences in C stocks, or that some process other than crop rotation is governing soil C dynamics across the farm.

3.3.4. Soil texture class

Soil texture class was evaluated for its influence on soil C stocks (Table 6). Silt loam soils had significantly higher C concentration (3.5 %) than channery silt loam, gravelly loam, or gravelly silt loam (3.0%), (2.9%), and (2.7 %), respectively. Mean bulk density for the different texture classes ranged from 1.14 to 1.21 g cm^{-3}, but only channery silt loam and gravely loam were significantly different from each other (Table 6). Silt loam soils had the highest C stocks (99.1 Mg C ha^{-1}); while there were no significant differences between any of the other texture classes.

The observed relationships between soil C and soil texture are consistent with Nichols, (1984) where SOC concentration was shown to have a strong positive correlation (r =.86) with clay content in Mollisol soils of the Sothern Great Plains. Similarly, Burke *et al*, (1989) found

that clay soils are less susceptible to C loss following cultivation, ant that C concentration increased with clay content. While we did not observe such strong relationships, silt loam (19 %) and channery silt loam (19 %) soils did tend to have higher SSURGO estimated clay content than gravely loam (12 %) and gravely silt loam (14 %) soils. Soil texture classes were unevenly distributed across the farm and showed high within-class variability, with gravely loam accounting for 66% of the samples and representing a range of values from 30 to 160 Mg C ha^{-1}. Due to this poor spatial coverage soil texture class will likely provide little benefit to sampling stratification or model prediction.

Table 6. Mean and coefficient of variation (CV) for C concentration (%C), C stocks, and bulk density (BD) for all major soil texture classes present on the farm. Mean values in each column with the same letter are not significantly different, α=0.05.

		C %		BD (g cm^{-3})		C-stock (Mg ha^{-1})	
	n	Mean	CV	mean	CV	mean	CV
Channery silt loam	56	3.0b	31	1.14b	16	90.1ab	23
Gravelly loam	213	2.9b	27	1.21a	13	89.8b	21
Gravelly silt loam	19	2.7b	23	1.18ab	12	86.0b	20
Silt loam	29	3.5a	32	1.15ab	16	99.1a	30
Entire farm	319	3.0	29	1.19	14	90.6	23

3.3.5. Manure application

Manure application rate was reclassified as a categorical variable with high, medium, low, pasture, and forest as categories (Table 1). Across these categories soil C concentration ranged from 2.7 to 3.4 %, with forest, pasture, and the high manure categories having significantly greater C concentration than either medium or low manure (Table 7). Land uses

with lower C concentrations tended to have higher bulk densities, with medium and low manure groups having significantly higher bulk densities, than forest, pasture, or high manure groups (Table 7). Mean soil C stocks for the different manure land use groups ranged from 87.9 Mg C ha^{-1} to 98.1 Mg C ha^{-1}. Soil C stocks for the high manure category (98.1 Mg C ha^{-1}), were significantly greater than both the low and medium categories (88.3 Mg C ha^{-1}) and (87.9 Mg C ha^{-1}), respectively; however, forest and pasture showed no significant differences in mean C stocks with any of the other manure categories (Table 7).

Positive relationships between soil C stocks and the cumulative amount of manure applied are consistent with finding by Sommerfeldt *et al*., (1988), where annual manure applications rates of 180 Mg ha^{-1} yr^{-1} (wet weight) were shown to increase soil organic matter concentration (SOM) by 0.57 % yr^{-1}. They found that after 11 yrs of applying cattle manure at rate of 180 Mg ha^{-1} yr^{-1}, SOM was increased by over 3%, and that tillage did not affect the rate of SOM accumulation. Similar findings by Jenkinson *et al*., (1990), showed that applying manure for 140 years at rate of 35 Mg ha^{-1} yr^{-1} increased C stocks from 27.5 Mg ha^{-1} to ~80 Mg ha^{-1} in the top 23 cm when continuous wheat was grown at the Rothamsted long-term field experiments. Crop land that was allowed to revert back to native vegetation for 140 years had similar C stocks as manure amended treatments, which is consistent with our findings. A recent meta-analysis by Maillard and Angers, (2013) found that the amount of applied manure-C was the dominant driver of increases in soil C stocks following manure application, and that manure application rate explained more variability in soil C stocks than animal species, land use, or soil properties. This may explain why we saw so few significant differences between the mean C stocks for different soil types, crop rotations or permanent land uses (Tables 2, 5, 6,7).

In concentrated animal feeding operations (CAFO) massive amounts of manure are generated annually, a single dairy cow may produce 46 liters of manure per day and a dairy farm in New York may have between 500 and 5,000 head of cattle. Operations of this scale must manage manure as a waste product, and often the main purpose of applying manure is waste disposal rather than meeting crop nutrient demands. For one field in our study site 2456 kg ha^{-1} of total N was applied over 4 growing seasons, this equates to 378 kl liquid manure ha^{-1} and 378 Mg solid manure ha^{-1} (Table 1). At this rate it is quite possible that C input from manure is equal to or even greater than C losses from tillage, which would explain why fields that have been planted with maize silage and moldboard plowed for the last five growing seasons, have soil C stocks equal to adjacent permanent pastures and forests. Manure application rate over 4 growing seasons seems to explain soil C spatial variability better than any of the other categorical predictors, and so it was used to define the stratified sampling approach.

Table 7. Mean and coefficient of variation (CV) for C concentration (%C), C stocks, and bulk density (BD) for land uses categorized by manure application rate. Mean values in each column with the same letter are not significantly different, α=0.05

		C %		BD (g cm^{-3})		C-stock (Mg ha^{-1})	
	n	mean	CV	mean	CV	mean	CV
Forest	20	3.4a	33	1.03c	20	91.6ab	27
Pasture	63	3.2a	30	1.11c	15	91.3ab	27
Crop-High Manure	52	3.3a	24	1.17b	13	98.1a	17
Crop- Med Manure	125	2.8b	29	1.23a	13	88.3b	24
Crop-Low-Manure	59	2.7b	19	1.24a	10	87.9b	16
Entire farm	319	3	29	1.19	14	90.6	23

3.4. Evaluation of sampling approaches

3.4.2. Representativeness of feature space

To test the effect of different sampling arrangements and densities on estimates of soil C stocks, the full data set was subsampled using random, stratified random and systematic removal. At the sampling intensity n=235, histograms for all sampling arrangements were very similar (Fig 7). Since the location of sample points for all sampling arrangements is derived from the initial full sampling grid n=319, it is not surprising that we see little difference between histograms at higher sampling densities. For the n=160 sampling density the stratified and systematic sampling approaches produce histograms that are in agreement with the full grid, with the random sampling arrangement producing a histogram that was slightly distorted.

At the lowest sampling density (n=83), the stratified approach resulted in a moderate distortion, suggesting that our manure stratification categories are not fully representing soil C stock feature space. The random sampling approach produced a substantially distorted histogram at n = 83 with high and low values over-represented. This is because sample clustering and poor spatial coverage resulted in an inaccurate representation of the distribution of soil C stocks across the farm (Fig. 1c). The random sampling approach for n = 83 was repeated five times to see how histograms varied for each randomization (Appendix B). For all randomizations the resulting histograms were not representative of the full sampling grid; however, there was some variation in the level of distortion. At n=83 the systematic sampling approach provided superior feature space coverage of soil C stocks, as can be seen by the nearly identical histograms produced by sample sizes of 319 and 83.

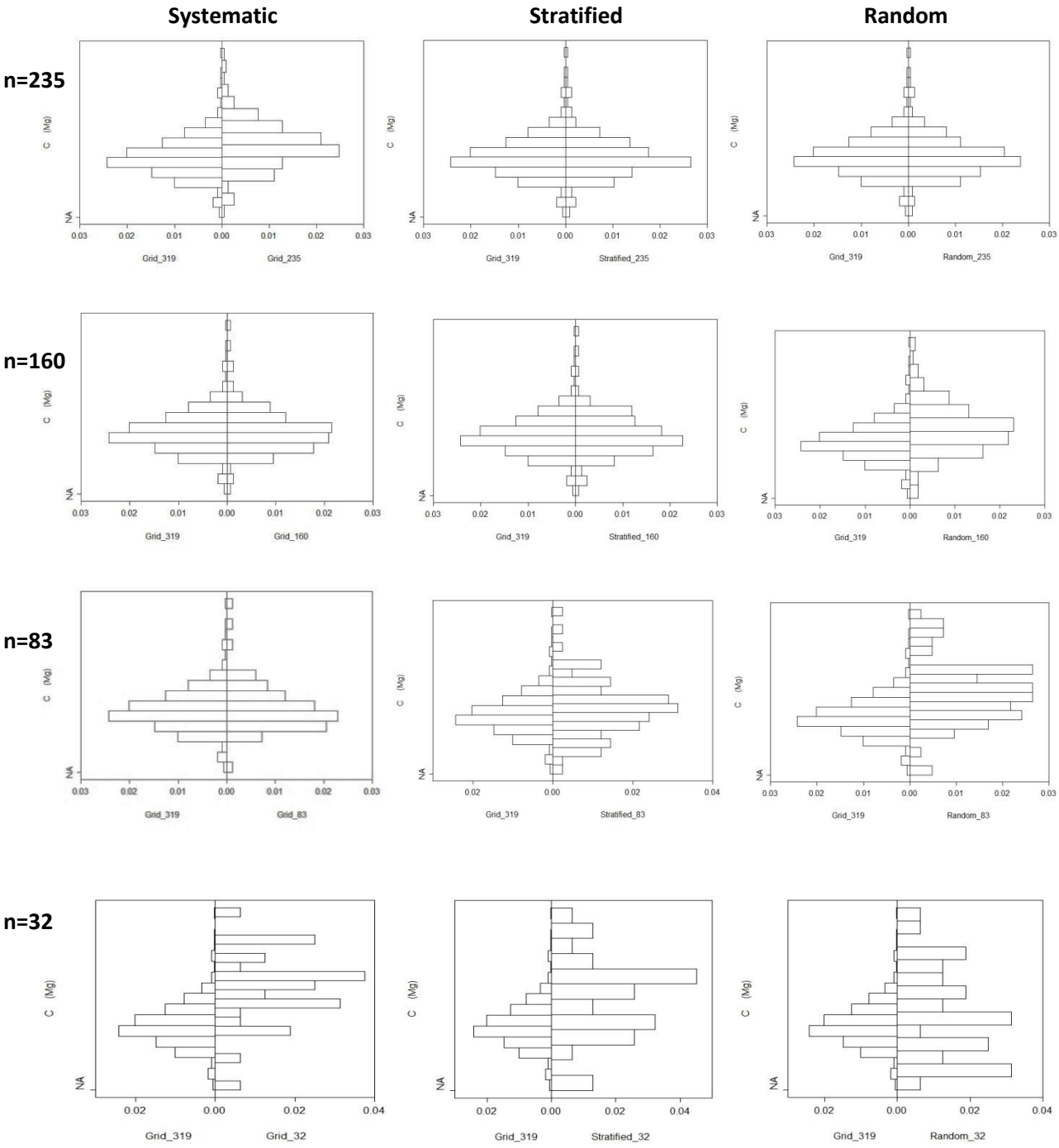


Fig. 7. Back to back histograms for different sampling densities and arrangements. The full sampling grid n=319 is on the left side of each histogram, while the reduced sampling arrangement is on the right side of each histogram.

When the sample size was further reduced to n=32 all sampling arrangements produced distorted histograms. These results indicate that a systematic grid with a sampling interval of 206 m will provide an adequate estimate of population parameters for soil C stocks on this farm.

3.4.3. Variogram modeling

Experimental variograms were computed for random (R83, R160, R235), stratified (M83, M160, M235), and systematic (S83, S160, S235) sampling arrangements at all sampling densities. Variograms were estimated for C concentration and fit by both MOM and REML methods. Variogram model parameters are summarized in Table 8. Variogram range parameters generally increased as sample size decreased, R235 had had the smallest range at 175 m and S83 had the largest range at 590 m. C concentration showed strong spatial structure with nugget to sill ratio less than 0.13 for all sampling arrangements except R83 where the nugget to sill ratios were equal to 0.37 and 0.57 for MOM and REML, respectively, and for S83 where the nugget to sill ratio was 0.49 for REML.

Table 8. Variogram model parameters (sill, nugget, range, and nugget to sill ratio (nug:sill)) for C concentration for three different sampling methods at three different sampling densities.

		MOM Semivariogram				REML Semivariogram			
	N	sill	nugget	range	nug:sill	sill	nugget	range	nug:sill
Systematic	319	0.73	0.08	206	0.10	0.79	0.08	220	0.11
	235	0.58	0.07	258	0.12	0.72	0.09	200	0.13
	160	0.70	0.04	251	0.06	0.90	0.06	250	0.06
	83	0.80	0.05	338	0.07	0.91	0.06	300	0.07
Random	235	0.64	0.06	175	0.09	0.86	0.08	175	0.10
	160	0.62	0.04	204	0.07	0.66	0.05	250	0.08
	83	0.63	0.23	485	0.37	0.68	0.39	450	0.57
Stratified	235	0.57	0.06	151	0.10	0.90	0.09	225	0.09
	160	0.69	0.06	178	0.08	0.81	0.08	225	0.09
	83	1.15	0.56	590	0.49	1.09	0.06	220	0.05

The range of the variogram describes the extent of spatial autocorrelation, or the maximum distance at which we observe trends in variation. To detect differences at this scale the sampling distance should be less than half the range of the variogram, thus the optimal sampling density is dependent on parameters of the variogram (Kerry and Oliver, 2003). Strong spatial structure (i.e. nug:sil < 0.25) will produce lower prediction errors when interpolated, regardless of the initial coefficient of variation (CV) of a data set (Kravchenko, 2003). This suggests that our confidence in an estimate may be more dependent on the spatial structure of the data rather than the variance, and that variables with a strong spatial structure will require a lower sample size for the same level of error (Bilgili et al., 2011).

3.4.4. Ordinary kriging predictions

C concentration was predicted by ordinary kriging using both MOM and REML variograms for all sampling arrangements, and the RMSE was determined by leave-one-out cross validation (Table 9). The systematic sampling approach produced the lowest RMSE for nearly all sample densities, but differences were most apparent at the lowest density where RMSE was equal to 0.699, 0.772, and 0.926 for S83, R83, and M83, respectively (Fig. 8). Equilateral triangular sample grids have been shown to produce more reliable variograms than other types of sampling arrangements,

Table 9. Root mean square error (RMSE) for C concentration predicted by ordinary kriging using both MOM and REML variograms. RMSE is provided for all sample sizes and arrangements.

		RMSE		
	n	Systematic	Stratified	Random
MOM	319	0.663	*	*
REML	319	0.675	*	*
MOM	235	0.672	0.666	0.765
REML	235	0.671	0.732	0.764
MOM	160	0.716	0.769	0.72
REML	160	0.706	0.772	0.732
MOM	83	0.719	0.926	0.772
REML	83	0.699	0.836	0.772

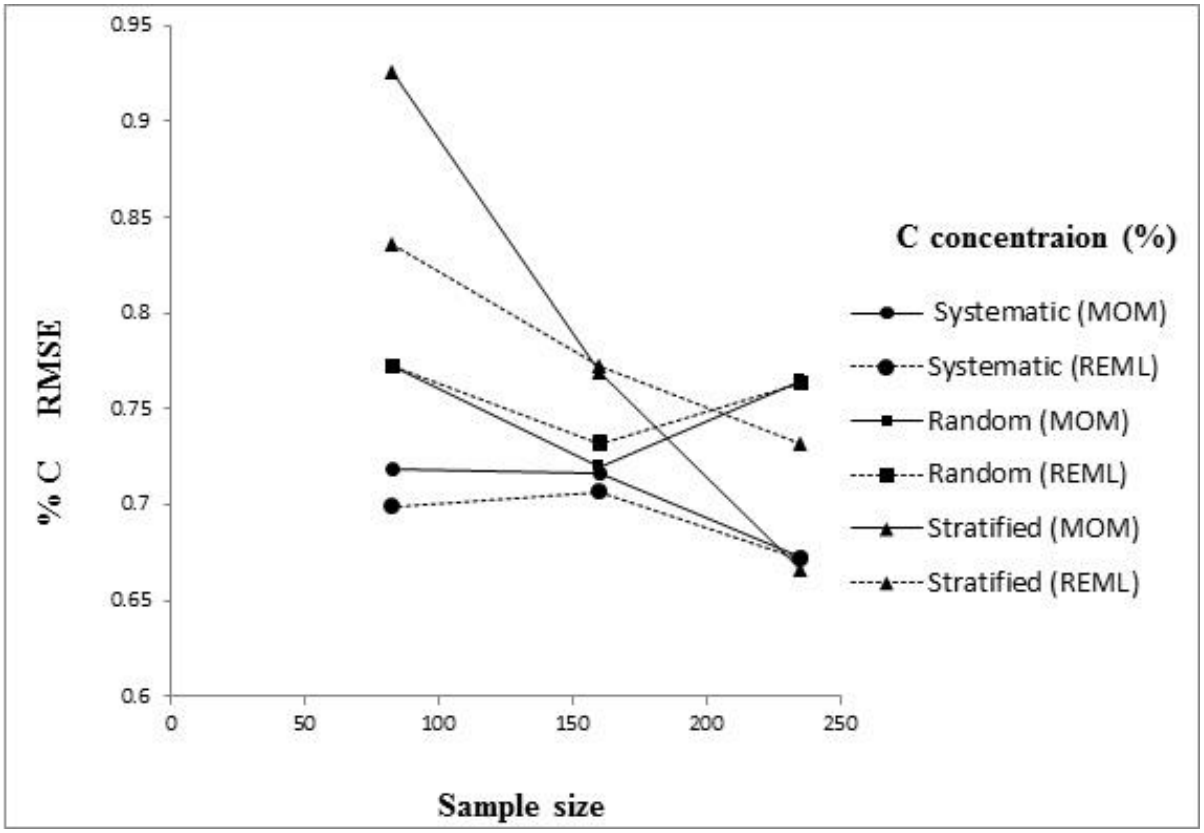


Figure 8. Root mean square error (RMSE) by sample number for different sampling arrangements using both MOM and REML variograms. RMSE was determined by ordinary kriging using leave one out cross validation.

which is supported by our results (Yfantis et al. 1987; McBratney et al. 1981). Additionally, the systematic approach ensured that the proportion and equal spacing of 1 m offset samples was held true for all sampling densities, providing a reasonable estimate of the nugget variance and lower prediction errors (De Gruijter, 2006). RMSE generally increased with decreasing sample size, except for R160 which had a slightly lower RMSE than R235, 0.764 compared to 0.732. For the stratified sampling approach RMSE increased at an almost linear rate from 0.666 to 0.926 as sample size declined, while for the systematic approach reducing the sample size from 160 to 83 resulted in a minimal decrease in RMSE from 0.706 to 0.699 (Fig. 6). Kerry and Oliver, (2007) found that REML variograms produced lower prediction errors than MOM

variograms for soil clay content when sample sizes where below 100; our findings corroborate this, except in the case of R83 where REML and MOM variograms produced identical errors

The systematic sampling approach appeared to represent the distribution of C stocks well for all sampling densities (Fig. 8), and so it is little surprise that it produced the lowest RMSE. However, it is unclear why RMSE increased so rapidly for the stratified approach. Stratified and random sampling arrangements presented in this study represent only a single realization of many possible sample arrangements; therefore we are unable to determine how different random arrangements would influence prediction results. Nevertheless, these results demonstrate the importance that sampling number and arrangement can have on prediction errors and the optimal sampling number.

3.5. Comparing design-based and model-based approaches

3.5.2 Comparing mean C stocks and uncertainty

Mean soil C stocks were estimated from the full sampling grid (n=319) and the lowest sampling density (n=83) for the entire farm using design-based and model-based approaches. The two methods were then compared based on the uncertainty of the estimates as determined by the RMSE (Table 10) (Chinembiri et al., 2013; Cambule et al., 2014). Overall, reducing sample size caused minor over-estimates. Design-based estimates consisted of the spatial mean (SM) where the mean was determined for each manure category and the entire farm, while model-based estimates were derived from the ordinary kriging (OK) predictions. For the full sampling grid mean soil C stocks for the entire farm were approximately equal for both approaches (90.6

and 90.8 Mg ha^{-1}), while at the lowest sampling density the design-based approach produced a slightly higher estimate (93.5 compared to 92.7 Mg ha^{-1}).

The design-based approach resulted in a higher RMSE for the entire farm at both sampling densities, 20.7 and 23.1 Mg ha^{-1} compared to 18.0 and 22.4 Mg ha^{-1} for the model-based approach. OK using the full sampling grid produced the lowest RMSE for nearly all manure categories, providing the most accurate estimate of mean soil C stocks. When estimating mean C stocks by OK using the full sampling grid, the high manure category had the greatest value at 94.5Mg ha^{-1}, followed by forest at 91.3 Mg ha^{-1}, then pasture at 89.8 Mg ha^{-1}, medium at 88.3 Mg ha^{-1}, and finally low at 87.9 Mg ha^{-1}.

When compared by the entire farm, differences between the two approaches are less extreme, 3.2 % diff for the design-based approach and 2.0 % diff for the model-based approach. By accounting for spatial auto correlation model-based approaches provided more reliable estimates of mean C stocks. Chinembiri *et al*., (2013) found that model-based approaches provided more accurate estimates of terrestrial C stocks in a forest watershed, and that design-based approaches tended to overestimate because predictions are all assigned the same weighting (Montes et al., 2005). We observed similar trends where design-based approaches produced higher estimates of total C stocks (Mg) for nearly every manure category and sampling density (Table 12).

Table 10**.** Mean C stocks and root mean square error (RMSE) for land uses categorized by manure application rate and for the entire farm. Estimates were generated by design-based (SM) and model-based (OK) approaches for two sample sizes (n = 319, n =83).

		Mean C stocks Mg ha^{-1}									
		n = 319					n = 83				
	Area		SM		OK			SM		OK	
	(ha)	n	Mean	RMSE	Mean	RMSE	n	Mean	RMSE	Mean	RMSE
Forest	15	20	91.6	25.1	91.3	23.4	9	102.1	17.4	88.5	17.8
Pasture	34	62	91.3	24.8	89.8	20.9	16	96.6	13.5	92.4	22.0
High Manure	39	51	98.1	16.6	94.5	18.2	16	93.7	16.6	94.4	15.0
Med Manure	81	125	88.3	21.4	88.6	16.9	27	90.8	30.6	93.6	25.3
Low Manure	46	59	87.9	13.9	89.1	14.2	16	83.7	22.7	88.9	25.8
Entire Farm	215	319	90.6	20.7	90.8	18.0	83	93.5	23.1	92.7	22.4

Table 11. Percent difference between estimates of mean C stocks for sample sizes of n = 319 and n =83 (% diff). Estimates were generated by design-based and model-based approaches for land uses categorized by manure application rate and for the entire farm.

	Mean C stocks Mg ha^{-1}					
	Design-based			Model-based		
	n=319	n=83	% diff	n=319	n=83	% diff
Forest	91.6	102.1	11.5	91.3	88.5	-3.1
Pasture	91.3	96.6	5.8	89.8	92.4	3.0
High Manure	98.1	93.7	-4.4	94.5	94.4	-0.2
Med Manure	88.3	90.8	2.8	88.6	93.6	5.5
Low Manure	87.9	83.7	-4.8	89.1	88.9	-0.3
Entire Farm	90.6	93.5	3.2	90.8	92.7	2.0

3.5.5. Comparing estimates of total C stocks

Total C stocks based on the full grid (n = 319) for the 232 ha study area were 19516, 18603, and 16217 Mg C for the design-based approach, the model-based approach, and SSURGO, respectively (Table 12). Once again we see that the model-based estimates for total soil C stocks vary less between the two sampling densities reinforcing our hypothesis that model-based approaches provide more reliable estimates for soil C stocks than design-based approaches. SSURGO produced the lowest estimates for all categories particularly underestimating the cropped categories. It is not surprising that SSURGO, based on measurements taken many years ago, would not accurately reflect recent land management effects of cropped systesm and therefore may tend to most closely represent relatively undisturbed pasture and forest soils.

Ordinary kriging with a sample size of 319 (OK319) produced the lowest RMSE for any estimate of mean soil C stocks for the entire farm and nearly all manure categories (Table 10).

As such we assume that the model-based approach provided the most accurate estimates of total soil C stocks, and evaluated the performance of all other methods by its percent difference with OK319 (Table 12). When the sampling density was reduced from 319 to 83 our model-based estimates of soil C stocks for the entire farm increased by only 2.4%, percent differences for the manure categories ranged from 0.2 to 5.5%. Design-based approaches tended to overestimate total C stocks, with a 7.8% increase for the entire farm. Percent difference with OK319 for the manure categories ranged from 4.7 to 19.7. SSURGO underestimated C stocks and had the largest % difference with OK319, with a 12.8% underestimate for the entire farm, and a difference ranging from 2.6 to 18.8% for the manure categories. However, the SSURGO estimates are of interest considering they came at no sampling or analytical costs. Gelder et al., (2011) showed that soil C concentration in top 15cm could be predicted at three field sites in Iowa by linear regression using SSURGO estimated C and surface reflectance from aerial imagery as predictors. Model R^2 and RMSE values ranged from 0.60 to 0.82 and 0.35 to 0.76%. SSURGO shows promise as a tool for estimating C stocks, this potential needs to be further investigated for different regions, soil types, and cropping systems.

Table 12. Total C stocks calculated for the entire farm and for each manure category using design-based (SM), model-based (OK), and SSURGO estimated approaches. Design-based and model-based estimates are provided for sample sizes of n = 319 and n = 83. Approaches are compared based the percent difference between OK using the full grid (% diff with OK n = 319) and all other estimates of total C stocks.

	Area	C stocks (Mg)					% diff with OK n=319			
			n=319		n=83			n=319	n=83	
	(ha)	SSURGO	SM	OK	SM	OK	SSURGO	SM	SM	OK
Forest	15	1187	1407	1218	1394	1181	-2.6	15.5	14.5	-3.1
Pasture	34	3490	4068	3591	3424	3698	-2.8	13.3	-4.7	3.0
High Manure	39	2572	3061	3168	3793	3163	-18.8	-3.4	19.7	-0.2
Med Manure	81	6001	7126	7122	7566	7517	-15.7	0.1	6.2	5.5
Low Manure	46	2967	3853	3503	3872	3492	-15.3	10.0	10.5	-0.3
Entire Farm	215	16217	19516	18603	20049	19050	-12.8	4.9	7.8	2.4

The specific objectives of a sampling campaign and the inherent biophysical characteristics of a site will ultimately dictate whether design-based or model-based approaches are most appropriate for soil C assessment. If high resolution estimates are needed then model-based approaches are usually best, but when estimates are needed for larger areas such as those defined by land uses, then the spatial structure of soil C stocks will dictate whether design-based or model-based approaches are most appropriate. At our study site, nearly all auxiliary variables displayed weak correlations with soil C stocks, and only manure categories showed significant difference between C stocks that could be logically explained by our knowledge of the system. Linear regression models incorporating auxiliary variables explained only a small amount of soil C variability (data not shown). Regression models incorporating elevation, slope, and manure categories produced the best predictions of soil C stocks with an R^2 of 0.16 (data not shown). We did not explore the benefits of regression kriging, because correlations were so weak, but it is possible the regression kriging would have resulted in decreased errors even with such low correlations with soil C.

4. Conclusions

We found that model-based approaches produced more reliable estimates for soil C stocks, with lower RMSE for nearly all manure land use categories. Additionally, when the sampling density was reduced from 319 to 83, model-based estimates showed a smaller percent difference between the two estimates, with mean and total soil C stocks for the entire farm differing by 2% from that of the full sampling grid (Table 11; Table 12). The systematic sampling arrangement was preferred over random or stratified arrangements because RMSE

increased little with the reduced sample size, and the distribution of soil C stocks for the lowest sampling density closely resembled the full sampling grid. REML fitted variograms generally produced a lower RMSE for OK predictions of soil C concentration when sample sizes were below 100, but for larger sample sizes both MoM and REML produced comparable results. Reductions in RMSE were not consistent across sampling arrangements; for the random sampling approach MoM and REML variograms produced similar errors for all sample sizes (Fig 8.). When assessing soil C stocks by a model-based approach we suggest that variograms be calculated by both MoM and REML so that the most reliable method can be determined.

SSURGO-based estimates consistently underestimated soil C stocks for the different landuses by 2.6 to 18.8 %, but given low cost (no sampling or analytical analyses) of this approach, may be of interest in some circumstances. Cropping history, soil series, and soil texture classes had little or no effect on C stocks. Forest, pasture and cropland had nearly identical mean C stocks, 91.6, 91.3, and 90.3, respectively. Manure application rate had a significant effect on C stocks, with the highest application rates being significantly greater than both the low and med application rates. Forest and pasture were not different from any of the manure categories. These trends imply that heavy manure application rates can maintain soil C stocks to levels equal to or greater than the relatively undisturbed native vegetation in this region. We hypothesize that cumulative manure application is controlling the magnitude of soil C stocks across cropland on this farm. Manure application records provide valuable information for estimating soil C stocks, future studies should evaluate the benefits of using manure application records for stratifying sampling, and modeling change.

5. Future Work

We were only able to obtain manure records for 4 growing seasons, and so could not estimate a longer-term cumulative amount of manure applied, which might have affected our manure categories. We plan to obtain manure records for at least 10 years, to compare the performance of short term records with longer records. Additionally, this study would benefit from a monte carlo simulation approach, where stratified and random sampling is repeated many times to determine the uncertainty of estimates associated with different sampling approaches. Adding regression kriging using auxiliary variables to these analyses would provide valuable information, particularly if longer term manure records can be obtained for the study site.

REFERENCES

Adams, W.A. 1973. The effect of organic matter on the bulk and true densities of some uncultivated podozolic soils. J. Soil Sci. 24(1): 10–17.

Allen, D.E., M.J. Pringle, K.L. Page, and R.C. Dalal. 2010. A review of sampling designs for the measurement of soil organic carbon in Australian grazing lands. Rangel. J. 32(2): 227–246.

Antle, J.M., S.M. Capalbo, S. Mooney, E.T. Elliott, and K.H. Paustian. 2002. A comparative examination of the efficiency of sequestering carbon in US agricultural soils. Am. J. Altern. Agric. 17(3): 109–115.

Antle, J., S. Capalbo, S. Mooney, E. Elliott, and K. Paustian. 2003. Spatial heterogeneity, contract design, and the efficiency of carbon sequestration policies for agriculture. J. Environ. Econ. Manage. 46(2): 231–250.

Bartholomeus, H.M., M.E. Schaepman, L. Kooistra, A. Stevens, W.B. Hoogmoed, and O.S.P. Spaargaren. 2008. Spectral reflectance based indices for soil organic carbon quantification. Geoderma 145(1–2): 28–36.

Bellamy, P.H., P.J. Loveland, R.I. Bradley, R.M. Lark, and G.J.D. Kirk. 2005. Carbon losses from all soils across England and Wales 1978-2003. Nature 437(7056): 245–248.

Bilgili, A.V., F. Akbas, and H.M. van Es. 2011. Combined use of hyperspectral VNIR reflectance spectroscopy and kriging to predict soil variables spatially. Precis. Agric. 12(3): 395–420.

Bishop, T.F.A., and A.B. McBratney. 2001. A comparison of prediction methods for the creation of field-extent soil property maps. Geoderma 103(1–2): 149–160.

Brus, D.J., and J.J. De Gruijter. 1997. Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with discussion). Geoderma 80(1): 1–44.

Brus, D.J., B. Kempen, and G.B.M. Heuvelink. 2011. Sampling for validation of digital soil maps. Eur. J. Soil Sci. 62(3): 394–407.

Burke, I.C., C.M. Yonker, W.J. Parton, C. V Cole, D.S. Schimel, and K. Flach. 1989. Texture, climate, and cultivation effects on soil organic matter content in US grassland soils. Soil Sci. Soc. Am. J. 53(3): 800–805.

Cambardella, C.A., T.B. Moorman, T.B. Parkin, D.L. Karlen, J.M. Novak, R.F. Turco, and A.E. Konopka. 1994. Field-scale variability of soil properties in central Iowa soils. Soil Sci. Soc. Am. J. 58(5): 1501–1511.

Cambule, A.H., D.G. Rossiter, and J.J. Stoorvogel. 2013. A methodology for digital soil mapping in poorly-accessible areas. Geoderma 192(0): 341–353.

Cambule, A.H., D.G. Rossiter, J.J. Stoorvogel, and E.M.A. Smaling. 2014. Soil organic carbon stocks in the Limpopo National Park, Mozambique: Amount, spatial distribution and uncertainty. Geoderma 213(0): 46–56.

Chatterjee, A., R. Lal, L. Wielopolski, M.Z. Martin, and M.H. Ebinger. 2009. Evaluation of different soil carbon determination methods. CRC. Crit. Rev. Plant Sci. 28: 164–178.

Chinembiri, T.S., M.C. Bronsveld, D.G. Rossiter, and T. Dube. 2013. The precision of C stock estimation in the Ludhikola watershed using model-based and design-based approaches. Nat. Resour. Res. 22(4): 297–309.

Conant, R.T., S.M. Ogle, E.A. Paul, and K. Paustian. 2011. Measuring and monitoring soil organic carbon stocks in agricultural lands for climate mitigation. Front. Ecol. Environ. 9: 169–173.

Conant, R.T., G.R. Smith, and K. Paustian. 2003. Spatial variability of soil carbon in forested and cultivated sites. J. Environ. Qual. 32(1): 278–286.

Cressie, N. 1990. The origins of kriging. Math. Geol. 22(3): 239–252.

Cressie, N.A.C. 1991. Statistics for Spatial Data. Wiley, New York.

Curtis, R.O., and B.W. Post. 1964. Estimating bulk density from organic-matter content in some Vermont forest soils. Soil Sci. Soc. Am. J. 28(2): 285–286.

Don, A., J. Schumacher, M. Scherer-Lorenzen, T. Scholten, and E.-D. Schulze. 2007. Spatial and vertical variation of soil carbon at two grassland sites—implications for measuring soil carbon stocks. Geoderma 141(3): 272–282.

Eswaran, H., E. Van Den Berg, and P. Reich. 1993. Organic carbon in soils of the world. Soil Sci. Soc. Am. J. 57(1): 192–194.

Gelder, B.K., R.P. Anex, T.C. Kaspar, T.J. Sauer, and D.L. Karlen. 2011. Estimating soil organic carbon in Central Iowa using aerial imagery and soil surveys. Soil Sci. Soc. Am. J. 75(5): 1821–1828.

Gregorich, E.G., C.F. Drury, and J.A. Baldock. 2001. Changes in soil carbon under long-term maize in monoculture and legume-based rotation. Can. J. Soil Sci. 81(1): 21–31.

De Gruijter, J.J. 2006. Sampling for natural resource monitoring. Springer, Berlin; New York.

Guo, L.B., and R.M. Gifford. 2002. Soil carbon stocks and land use change: a meta analysis. Glob. Chang. Biol. 8(4): 345–360.

Hengl, T., G.B.M. Heuvelink, and D.G. Rossiter. 2007. About regression-kriging: From equations to case studies. Comput. Geosci. 33(10): 1301–1315.

Hengl, T., G.B.M. Heuvelink, and A. Stein. 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120(1–2): 75–93.

Holmes, K.W., A. Wherrett, A. Keating, and D. V Murphy. 2012. Meeting bulk density sampling requirements efficiently to estimate soil carbon stocks. Soil Res. 49 (8): 680–695.

Hudson, B.D. 1994. Soil organic matter and available water capacity. J. Soil Water Conserv. 49(2): 189–194.

Jenkinson, D.S., S.P.S. Andrew, J.M. Lynch, M.J. Goss, and P.B. Tinker. 1990. The turnover of organic carbon and nitrogen in soil [and discussion]. Philos. Trans. R. Soc. London. Ser. B Biol. Sci. 329(1255): 361–368.

Kerry, R., and M. Oliver. 2003. Variograms of ancillary data to aid sampling for soil surveys. Precis. Agric. 4(3): 261–278.

Kerry, R., and M.A. Oliver. 2007. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. Geoderma 140(4): 383–396.

Knotters, M., D.J. Brus, and J.H. Oude Voshaar. 1995. A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. Geoderma 67(3–4): 227–246.

Kravchenko, A.N. 2003. Influence of spatial structure on accuracy of interpolation methods. Soil Sci. Soc. Am. J. 67(5): 1564–1571.

Kravchenko, A.N., and G.P. Robertson. 2007. Can topographical and yield data substantially improve total soil carbon mapping by regression kriging? Agron. J. 99(1): 12–17.

Kravchenko, A.N., G.P. Robertson, S.S. Snap, and A.J.M. Smucker. 2006. Using information about spatial variability to improve estimates of total soil carbon. Agron. J. 98(3): 823–829.

Lal, R. (Ed). Assessment Methods for Soil Carbon. Boca Raton, Fla. : Lewis Publishers.

Lal, R. 2004. Soil carbon sequestration impacts on global climate change and food security. Science (80-.). 304(5677): 1623–1627.

Lark, R.M. 2000. Estimating variograms of soil properties by the method-of-moments and maximum likelihood. Eur. J. Soil Sci. 51(4): 717–728.

Lark, R.M. 2009. Estimating the regional mean status and change of soil properties: two distinct objectives for soil survey. Eur. J. Soil Sci. 60(5): 748–756.

Madari, B.E., J.B. Reeves III, M.R. Coelho, P.L.O.A. Machado, H. De Polli, R.M. Coelho, V.M. Benites, L.F. Souza, and G.W. McCarty. 2005. Mid and near infrared spectroscopic determination of carbon in a diverse set of soils from the brazilian national soil collection. Spectrosc. Lett. 38(6): 721–740.

Maillard, É., and D.A. Angers. 2013. Animal manure application and soil organic carbon stocks: a meta-analysis. Glob. Chang. Biol.: n/a–n/a.

Matheron, G. 1962. Trait de geostatistique applique 1. 1. de. Technip, Paris.

McBratney, A.B., M. de L. Mendonça Santos, and B. Minasny. 2003. On digital soil mapping. Geoderma 117(1): 3–52.

McCarty, G.W., J.B. Reeves, V.B. Reeves, R.F. Follett, and J.M. Kimble. 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. Soil Sci. Soc. Am. J. 66(2): 640–646.

Minasny, B., A.B. McBratney, B.P. Malone, and I. Wheeler. 2013. Digital mapping of soil carbon. Adv. Agron. 118: 1–47.

Montes, F., M.J. Hernández, and I. Cañellas. 2005. A geostatistical approach to cork production sampling estimation in Quercus suber forests. Can. J. For. Res. 35(12): 2787–2796.

Mooney, S., K. Gerow, J. Antle, S. Capalbo, and K. Paustian. 2007. Reducing standard errors by incorporating spatial autocorrelation into a measurement scheme for soil carbon credits. Clim. Change 80(1-2): 55–72.

Nichols, J.D. 1984. Relation of organic carbon to soil properties and climate in the southern Great Plains. Soil Sci. Soc. Am. J. 48(6): 1382–1384.

Ogle, S.M., F.J. Breidt, and K. Paustian. 2005. Agricultural management impacts on soil organic carbon storage under moist and dry climatic conditions of temperate and tropical regions. Biogeochemistry 72(1): 87–121.

Pautsch, G.R., L.A. Kurkalova, B.A. Babcock, and C.L. Kling. 2001. The efficiency of sequestering carbon in agricultural soils. Contemp. Econ. Policy 19(2): 123–134.

Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30(7): 683–691.

Post, W.M., and K.C. Kwon. 2000. Soil carbon sequestration and land-use change: processes and potential. Glob. Chang. Biol. 6(3): 317–327.

Pringle, M.J., D.E. Allen, R.C. Dalal, J.E. Payne, D.G. Mayer, P. O'Reagain, and B.P. Marchant. 2011. Soil carbon stock in the tropical rangelands of Australia: Effects of soil type and

grazing pressure, and determination of sampling requirement. Geoderma 167–168(0): 261–273.

Puget, P., and R. Lal. 2005. Soil organic carbon and nitrogen in a Mollisol in central Ohio as affected by tillage and land use. Soil Tillage Res. 80(1): 201–213.

Reeves III, J.B. 2010. Near-versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done? Geoderma 158(1): 3–14.

Reeves III, J.B., R.F. Follett, G.W. McCarty, and J.M. Kimble. 2006. Can near or mid-infrared diffuse reflectance spectroscopy be used to determine soil carbon pools? Commun. Soil Sci. Plant Anal. 37(15-20): 2307–2325.

Sherrod, L.A., G. Dunn, G.A. Peterson, and R.L. Kolberg. 2002. Inorganic carbon analysis by modified pressure-calcimeter method. Soil Sci. Soc. Am. J. 66(1): 299–305.

Simbahan, G.C., and A. Dobermann. 2006. Sampling optimization based on secondary information and its utilization in soil carbon mapping. Geoderma 133(3): 345–362.

Simbahan, G.C., A. Dobermann, P. Goovaerts, J. Ping, and M.L. Haddix. 2006. Fine-resolution mapping of soil organic carbon based on multivariate secondary data. Geoderma 132(3): 471–489.

Singh, K., B.W. Murphy, and B.P. Marchant. 2013. Towards cost-effective estimation of soil carbon stocks at the field scale. Soil Res. 50 (8): 672–684.

Sommerfeldt, T.G., C. Chang, and T. Entz. 1988. Long-term annual manure applications increase soil organic matter and nitrogen, and decrease carbon to nitrogen ratio. Soil Sci. Soc. Am. J. 52(6): 1668–1672.

Syswerda, S.P., A.T. Corbin, D.L. Mokma, A.N. Kravchenko, and G.P. Robertson. 2011. Agricultural management and soil carbon storage in surface vs. deep layers. Soil Sci. Soc. Am. J. 75(1): 92–101.

Team, R.D.C. 2005. R: A language and environment for statistical computing. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2013. url: http://www. R-project. org.

Viscarra Rossel, R.A., D.J.J. Walvoort, A.B. McBratney, L.J. Janik, and J.O. Skjemstad. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131(1): 59–75.

Webster, R., and M.A. Oliver. 1992. Sample adequately to estimate variograms of soil properties. J. Soil Sci. 43(1): 177–192.

Webster, R Oliver, M. A.,, R. 2001. Geostatistics for Environmental Scientists. John Wiley & Sons, Chichester, England; New York.

West, T.O., and W.M. Post. 2002. Soil organic carbon sequestration rates by tillage and crop rotation. Soil Sci. Soc. Am. J. 66(6): 1930–1946.

West, L.T., S.W. Waltman, S. Wills, T.G. Reinsch, E.C. Benham, C.S. Smith, and R. Ferguson. 2010. Soil carbon stocks in the US: Current data and future inventories. p. 14. *In* Proceedings of International Workshop on Evaluation and Sustainable Management of Soil Carbon Sequestration in Asian Countries. Bogor, Indonesia.

Worsham, L., D. Markewitz, and N. Nibbelink. 2010. Incorporating spatial dependence into estimates of soil carbon contents under different land covers. Soil Sci. Soc. Am. J. 74(2): 635–646.

Xiao, X., G. Gertner, G. Wang, and A. Anderson. 2005. Optimal sampling scheme for estimation landscape mapping of vegetation cover. Landsc. Ecol. 20(4): 375–387.

Zhang, S., Y. Huang, C. Shen, H. Ye, and Y. Du. 2012. Spatial prediction of soil organic matter using terrain indices and categorical variables as auxiliary information. Geoderma 171–172(0): 35–43.

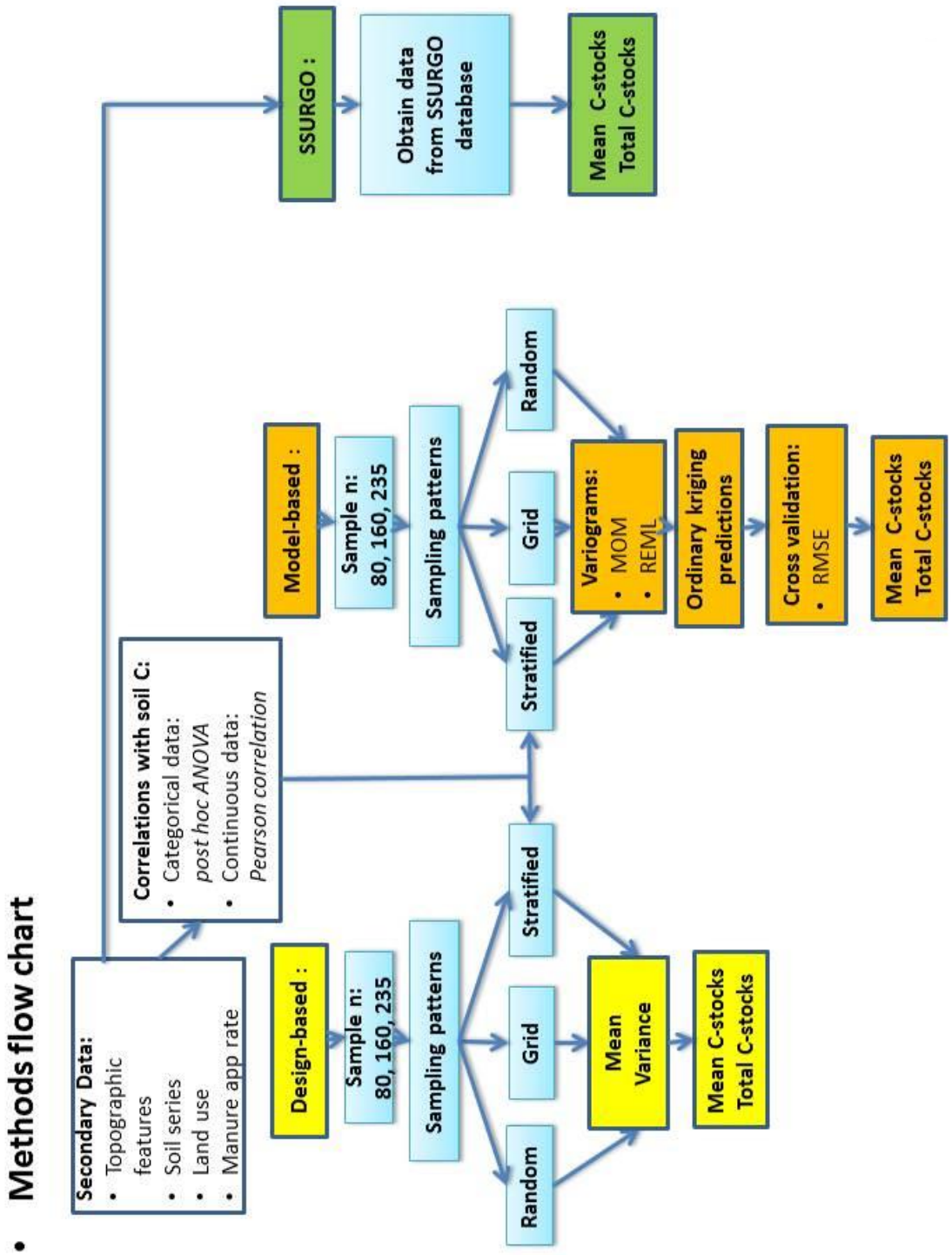


Figure 1. Methodological flow chart for comparing design-based, model-based, and SSURGO estimated approaches to soil C assessment.

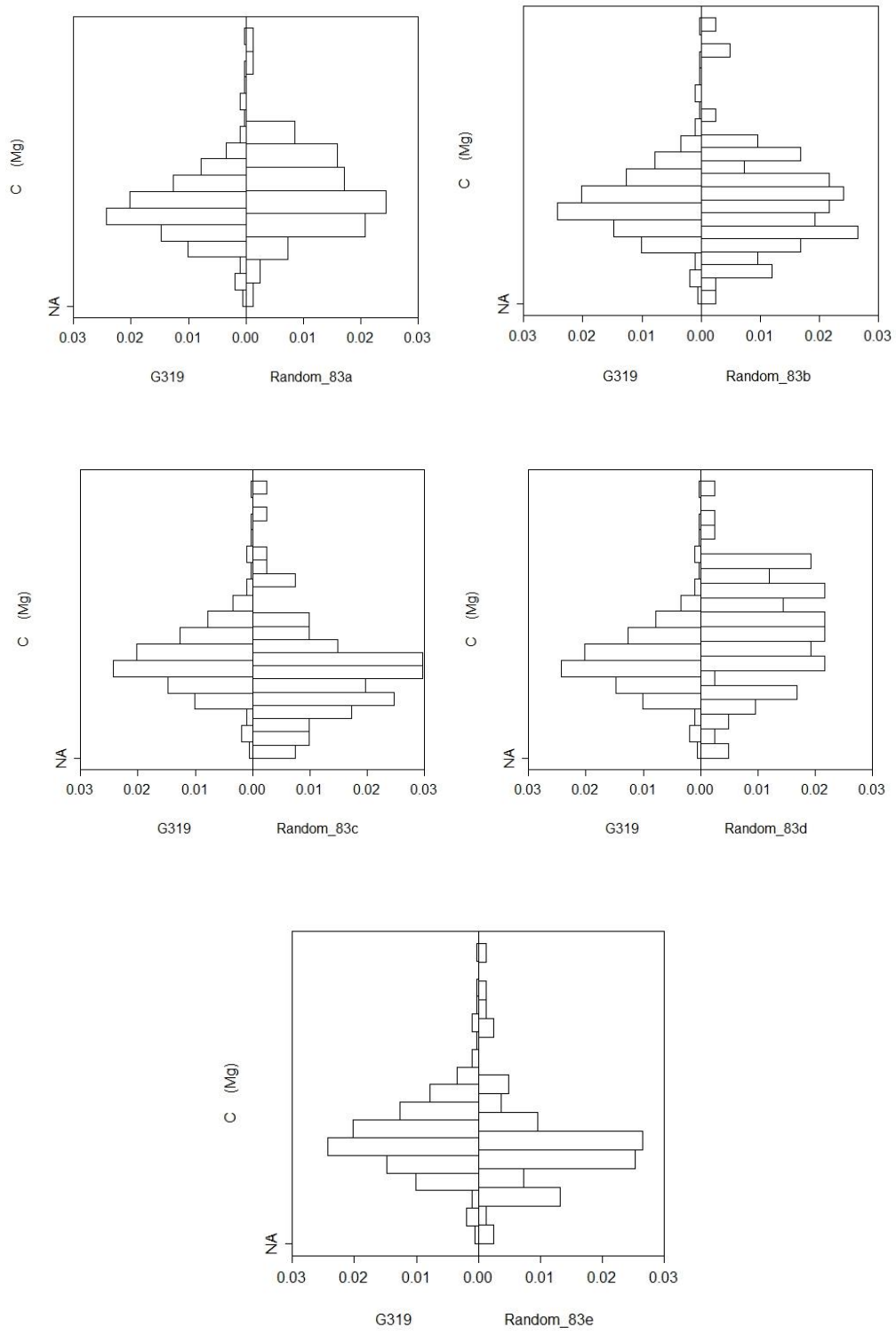


Figure 2. Back to back histograms for five different randomizations of the random sampling arrangement at n = 83. The full sampling grid n = 319 is on the left side of each histogram, while the reduced sampling arrangement is on the right side of each histogram.